SELF-ASSEMBLY OF PROTEIN DERIVED PEPTIDES

A thesis submitted in partial fulfilment of the requirements for the

Degree of Doctor of Philosophy in Biochemistry

2016

Rishi Kumar Pandey

University of Canterbury

Acknowledgements

A lot of hard work has gone into the preparation of this thesis, and a lot of people need to be thanked for their constant support and guidance. First and foremost I would like to thank my supervisor Prof. Juliet Gerrard for giving me the opportunity to come to New Zealand and work on the project and for her constant support, guidance and encouragement. My sincere thank you goes to co-supervisor Dr. Celine Valéry (RMIT University, Australia) for her confidence in my abilities, her guidance throughout the project, her patience, criticism, knowledge and not giving up on me at my lowest. I am especially thankful to you for helping me to understand the principles of transmission electron microscopy and the technique, which helped me greatly in this project.

I am very grateful and thankful to my supervisor Dr. Grant Pearce for having faith in me, his patience, expert advice and his critical feedback. I am sincerely thankful to my co-supervisors, Dr. Renwick Dobson and Dr. Jolon Dyer for their regular guidance and inputs which helped to determine the directions of this project. A big thank you goes to Dr. Ali Reza Nazmi, Dr. Amy Phillips, Dr. Hironori Suzuki, Dr. Luigi Sasso and Dr. Moritz Lassé, for their regular advice. All this work would not have been possible without our lab manager Jackie Healey and thank you for ensuring that the lab had a constant supply of everything and for regular trips to Lincoln for TEM. A special thank you goes to the Riddet Institute (Palmerston North, New Zealand), AgResearch (Lincoln, New Zealand), MacDiarmid Institute and the Biomolecular Interaction Centre for funding my PhD. research. I would also like to thank you, James Vernon (AgResearch, Lincoln), for giving me regular time on TEM.Also, you have always been fun to work with.

A special thank you goes to all my fellow colleagues at UC, Amy Y, Alexandra, Akshita, Ayelen, Arvind, Deepti, Helen, Jen, Jeremy, Eric, Kat, Manmeet, Nivas, and Rachel for their enjoyable discussions and providing me good company in the lab. I would also like to thank you Ayelen, Robert, Stephanie and Wafaa for the regular lunch trips to the undercroft and fun-filled conversations both in and outside the lab.

My especial thank you also goes to friends outside the university, that made Christchurch very much like home. Akshita, Amol, Arohi, Chandan, Mohit, Shishir, Sitara, Tanmay and Utkarsh, it would have been really difficult had you all not been there for me as constant pillars of support. Thank you also, must go to Joe, Noeline and Jesse for all the fun times and pushing me to give my best. A special thank you goes to Anu, Akshay and Shruti for being there for me in my good and bad times. This would not have been possible without you all.

Last but not the least I would like to thank my parents, sister, and brother back home. I have missed you greatly and could not have finished this journey without your encouragement. Thank you for your love, affection, support and your abiding faith in me.

Table of contents

Table of contents

Abbreviations

Å	angstrom
A_{280}	absorbance at 280 nm
Amp	ampicillin
*Ap*Tpx	Aeropyrum pernix peroxiredoxin
ATR	attenuated total reflectance
AUC	analytical ultracentrifugation
AU	absorbance units
BSA	bovine serum albumin
β-Lac	β-lactoglobulin
°C	degree Celsius
CD	circular dichroism
Chl	chloramphenicol
CV	column volumes
cryo-EM	cryo-electron microscopy
Cys	cysteine
d	*derivative*
dt/dc	*ThT fluorescence intensity per mg of peptide*
Da	kilo Daltons
DAPDC	*Mycobacterium tuberculosis* diaminopemilate decarboxylase
DLS	dynamic light scattering
DNA	deoxyribonucleic acid
DTT	dithiothreitol
E. coli	*Escherichia coli*
ε_{280}	extinction coefficient
FTIR	Fourier transform infrared spectroscopy
g	grams
h	hours
HEPES	N-2-hydroxyethlpiperazine-N'-2-ethane sulphonic acid
HFIP	hexaflouro-2-propanol
His	histidine
*h*Prx3	human peroxiredoxin 3

I(q)	*q* dependent intensity
IMAC	immobilized metal affinity chromatography
IPTG	isopropyl β-D-1-thiogalactopyranoside
kDa	kiloDaltons
L	litre
LB	Luria Bertani
Lsm	Sm-like
µg	micrograms
µL	microliter
µM	micromolar
µm	micrometre
mL	millilitre
mM	millimolar
mins	minutes
MAS	Magic angle spinning
MES	2-(N-morpholino)ethanesulfonic acid
Lsmα	*M. thermoautotrophicum* Lsm
MW	molecular weight
MWCO	molecular weight cut off
nm	nanometres
nM	nanomolar
NMR	nuclear magnetic resonance
OD_{600}	optical density at 600 nm
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffer saline
PDB	protein data bank
Prx	peroxiredoxin
psi	pounds per square inch
r	recombinant
RNA	ribonucleic acid
rpm	revolutions per minute
s	seconds
q	scattering vector
S200	Superdex 200 matrix

SAXS	small angle X-ray scattering
SDS	sodium dodecyl sulfate
SEC	size exclusion chromatography
SLS	static light scattering
SOB	super optimal broth with glucose added
TEM	transmission electron microscopy
TFA	triflouroacetic acid
ThT	Thioflavin-T
TMV	tobacco mosaic virus
TRAP	tRNA attenuation protein
Tris	tris(hydroxymethyl)aminomethane
U	unit
UV	ultraviolet
w/v	unit weight per unit volume
w/w	unit weight per unit weight
WAXS	wide angle X-ray scattering

List of figures

List of tables

This thesis has contributed to the following publications:

Valery, C., **Pandey, R**. & Gerrard, J. A. (2013). Protein β-interfaces as a generic source of native peptide tectons. *Chemical Communications* **49**, 2825-2827.

Eakins, G. L., **Pandey, R**., Wojciechowski, J. P., Zheng, H. Y., Webb, J. E. A., Valéry, C., Thordarson, P., Plank, N. O. V., Gerrard, J. A. & Hodgkiss, J. M. (2015). Functional Organic Semiconductors Assembled via Natural Aggregating Peptides. *Advanced Functional Materials* **25**, 5640-5649.

Abstract

Self-assembling building blocks have become of increasing interest in the field of bionanotechnology due to their ability to self-assemble into defined geometrical shapes. Nature is abundant with examples of functional biological assemblies that are either rich in β-sheets or coiled coils. This research work investigates peptide sequences derived from simple protein interfaces, as molecular tectons for use in bionanotechnology. Protein interface sequences were chosen as a design source of peptide tectons, as they are naturally optimized to drive self-assembly in a highly controlled and regulated manner. This thesis work focused primarily on the simple β-continuous and the helical coiled coil protein interfaces.

The peptide sequences designed from different protein-β interfaces all self-assembled into well-ordered nanostructures that were β-sheet rich and exhibited liquid crystallinity. SAXS and FTIR confirmed that the extended peptide nanostructures have the common architecture of β-sheets organized in a repetitive manner over long distances. The self-assembly of the peptide designed from the homodimer interface of *bovine* Prx3 was studied by thoughtfully altering appropriate residue in the peptide sequence. The peptide sequence was tolerant to the different variations introduced, with variant peptide β-sheet assemblies judged by ThT binding assays and SAXS consistent with the parent peptide. The self-assembling peptide sequences were used to template micron long nanotubes of the *h*Prx3 protein. Peptides designed from the coiled-coil interface of *ovine* keratin protein self-assembled into extended filament-like nanostructures with dimensions consistent with that of full-length keratin IFs.

The results presented in this thesis work aid in understanding the natural self-assembly process through *in vitro* self-assembly of peptide mimics. The self-assembling peptide sequences were used to produce extended nanotubes and nanosheets of *h*Prx3 protein. Future work is needed to show the functionalization of such extended peptide and protein structures, though, this work showcases the potential of these peptide sequences as building blocks for hybrid biomaterials.

1 Chapter one

Introduction

1.1 Context

Self-assembly is a process wherein building blocks orient and organize themselves in an energetically favored assembly, both in synthetic and biological systems[1]. Self-assembling peptides are a burgeoning area of attracting interest from both a fundamental viewpoint and due to their many potential applications in biomedical sciences such as tissue engineering and drug delivery as well as biosensors, conductive nanowires and nanocircuit fabrication in nanotechnology[2,3]. Tectons are simple building blocks that assemble into well-ordered supramolecular structures through self-assembly[1]. This research explores the potential of small peptide sequences inspired from self-assembling proteins as molecular tecton. This chapter presents an overview of the current understanding of short self-assembling peptide systems and an introduction to novel sources of short sequence design.

1.2 Molecular interactions in self-assembly

The interactions governing self-assembly of biomolecules in biology are mainly non-covalent forces such as electrostatic interactions, hydrophobic interactions, hydrogen bonding, van der Waals' interactions, aromatic stacking, and metal coordination. Molecular self-assembly is a ubiquitous process in biological systems and underlies the formation of a range of complex, well-defined biological structures joined by non-covalent bonds. Non-covalent interactive forces are weak individually but in synergy and higher number can generate stable and ordered assemblies[4] (Figure 1.1).

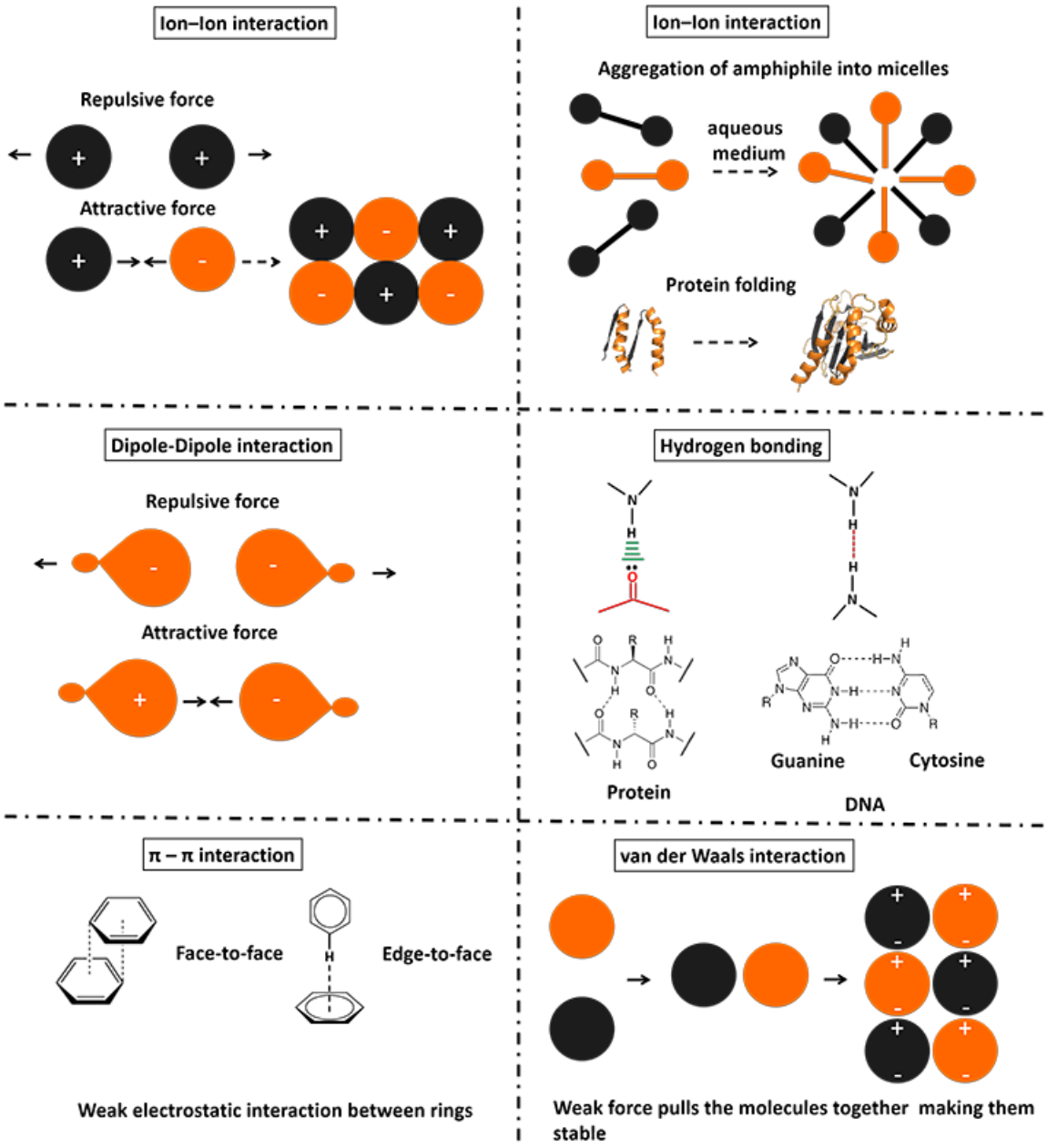


Figure 1.1: *Illustration of the non-covalent interactions involved in molecular self-assembly*[5]*. Reproduced with permission from Mendes et al. (2013)*[5]*.*

Electrostatic or charge interactions are Columbic forces that arise due to attraction or repulsion between dissimilar or similar charged entities. In proteins, ionic self-assembly is a direct result of the charge interactions between side chains of charged amino acid residues, furthermore also between terminal ammonium and carboxylate groups of the peptide chain. This leads to strong and long range forces in peptides and proteins[6].

Hydrogen bonds are the primary driver of peptide self-assembly, resulting from the association of an electronegative atom with a hydrogen atom attached to a relatively electronegative atom. Usually, the electronegative atoms are nitrogen (N) and oxygen (O) in natural systems and fluorine (F) in synthetic molecules. Hydrogen bonding forces are strong and directional (short or long) intermolecular bonds with high bond strength (10-40 kJ/mol)[6]. A salt bridge is a combinatorial force arising as a result of a combination of hydrogen bond and electrostatic interactions formed between the anionic carboxylate ($RCOO^-$) and the cationic ammonium (RNH_3^+) or guanidinium [$(RNHC(NH_2)_2)^+$] [6].

Hydrophobic effects arise due to the repulsion of the solute by the solvent. Non-polar molecules and motifs tend to come together to avoid aqueous surroundings. Hydrophobic interactions occur via side chain interactions between complementary non-polar amino acid residues. They are key driving forces in self-assembly of peptide amphiphiles[7].

van der Waals' forces are comparatively weak repulsive or attractive forces. They arise from a dipole or induced dipole interactions at the atomic and molecular level, due to continuous oscillation of various atoms of the side chain within a peptide chain[8]. Though weak individually, several van der Waals' forces collectively have a significant effect, they influence tertiary structure of proteins[9] and higher order peptide assemblies[10].

Aromatic stacking known as π-π stacking is an attractive interaction between aromatic rings in close contact oriented face to face like two heads of a coin[11]. It is an important non-covalent interaction reported in peptide motifs with short continuous stretches of single aromatic amino acid residues.

1.3 Self-assembly in synthetic systems

The spontaneous organization of individual components into well-ordered supramolecular structures is the key feature of self-assembly. Self-assembly in this context is the association of two or more chemical species held together by intermolecular forces. Strategies for building synthetic building blocks that can self-assemble in a highly efficient and selective manner have relied on manipulation of the different non-covalent interactions within a molecular architecture[12]. Early design principles for synthetic self-assembly focused on molecular recognition of a guest molecule (substrate) by a host molecule (receptor). The organization of the receptor molecule with intermolecular cavities around the guest molecule through numerous non-covalent interactions yielded an inclusion complex, called a "crypt"[12]. An array of architectures of different sizes and shapes were created by manipulation of the receptor chemistry and stereochemistry of non-covalent interactions. An important example is the Nobel Prize winning work by Lehn, Cram, and Pedersen on cryptands (a family of polycyclic multidentate ligands that bind to substrates in a crypt). Lehn & co-workers designed alkali metal cryptates inspired from the structure of the cyclopeptide valinomycin that mediates potassium ion transport in mitochondria. The design model comprises a three-dimensional interior cavity of macrocyclic polyethers (crown ethers), creating a binding site for the guest ions[13].

Diverse set of cryptands has been developed using the design model including macro-bicyclic cryptates selective for cations (Li^+, Na^+ and K^+) and anionic cryptates from protonated polyamines selective for the spherical halide ions F^- and Cl^-. Manipulation of the macro-polycyclic ether allowed for control of the recognition site and overall size of the crypt, such as the cryptate with a tetrahedral recognition site selective for the NH_4^+ cation[14,15]. The host-guest chemistry was further explored through molecular recognition, multivalent ligands, and

molecular hosts. An important example of self-assembly through molecular recognition is the bowl-shaped ligand structure named "carcerand" that interacts specifically with the guest molecules like a receptor[16]. Aoyama and co-workers described different assemblies of the carcerand resorcin[4]arenes with diacids, alcohols, and steroids as guests in organic solvents held stably through intermolecular hydrogen bonding[17,18]. Figure 1.2 shows some important examples of supramolecular structures created from host-guest chemistry.

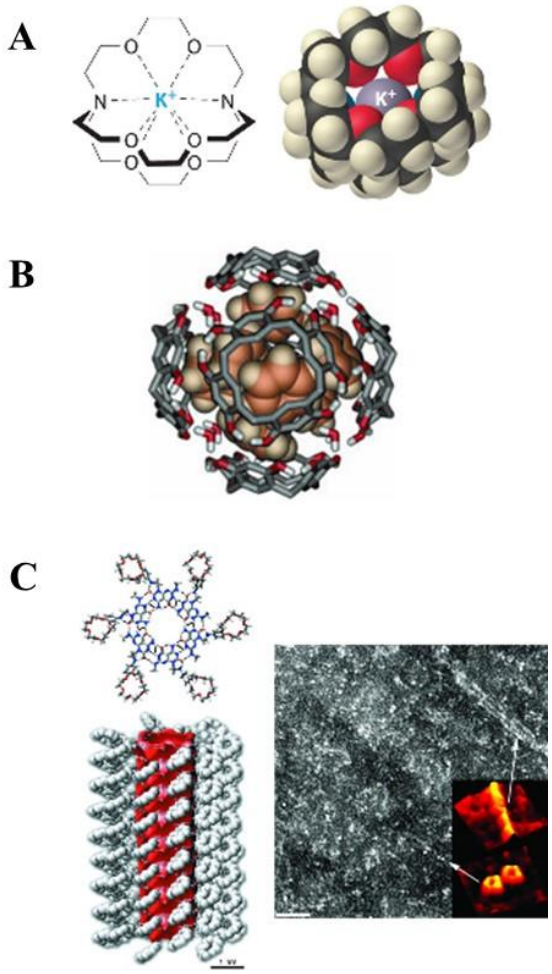


Figure 1.2: *Examples of synthetic self-assembled architectures: (A) 3-dimensional structure of the potassium complex of a 2,2,2-cryptand[19]. (B) Hexameric capsules from the self-assembly of resorcin[4]arenes[20]. (C) The structure of the helical rosette nanotubes that result from the entropically driven self-assembly of synthetic G^C motifs in water. The TEM micrographs showing the 2-dimensional view of the hollow nanotubes[21]. Taken with permission from Vicens et al. (2006)[19], Evan-Salem et al. (2006)[20] and Fenniri et. al. (2002)[21].*

Multivalent ligands are another class of ligand complexes that are composed of recognition elements (REs) conjugated to a core scaffold. Ligand Complexes with controlled size and stoichiometry have been created that resemble natural protein-protein receptor-substrate complexes. An important example includes the β-cyclodextrin grafted with hyaluronic acid (HA-β-CD) multivalent binding to self-assembled monolayers (SAMs) functionalized with ferrocene (SAM-Fc), that resembles the multivalent binding of hyaluronic acid to cell surfaces. The system shows potential application as clinical products for cell therapy in biomedical science[22].

Molecular hosts are an important example of polytopic receptor molecules containing two or more discrete binding subunits for simultaneous complexation of identical or different guest molecules. This design has allowed for the production of tailor-made polymer assemblies in water; an example is the cucurbit[8]uril (CB[8]) molecular host that was used to create a redox-sensitive molecular loop lock complex[23]. Tailored chemical motifs tuned to self-assemble into oligomer complexes of specific lengths include the rosette tubes with an open central channel of 1.1 nm diameter[21] and analogs of the guanosine tetrad motifs[24].

Amphiphilic copolymer and amphiphilic diblock copolymers are another important class of synthetic building blocks that can self-assemble into vesicles and micelles[25]. This model has been successfully exploited to produce pH-responsive multivesicular structures that mimic natural transmembrane channels[26]. The ability to tune self-assembly of building blocks enables the creation of a range of different supramolecular structures with desired properties.

1.4 Self-assembly in biological systems

Biological systems have evolved to utilize building blocks naturally tuned to work in synergy with each other and their environment to produce well-ordered structures with functionality. The next paragraphs present a few examples of some biological systems.

Nucleic acids:

Deoxiribonucleic acid (DNA) that carries the genetic information is an important example of self-assembled architecture. DNA architecture comprises of four different nucleotide building blocks organized into the thermostable double helical DNA complex stabilized through intermolecular hydrogen bonding[27].

Lipids:

Lipids are essential biomolecules that show variety in structure and are important structural components of the cell membranes. Amphipathic lipids self-assemble *in vitro* into different structures, the most common architectures encountered being the bilayer arrangement as found in the cell membrane[28], micelle structure, and vesicles[29].

Carbohydrates:

Carbohydrates have a rich topological portfolio and make up a huge repertoire of structures that are part of the exoskeleton and cell wall[30].

Proteins:

Proteins are complex macromolecules made up of 20 different amino which defines by their amino acid sequence which controls their secondary structure conformation and the complex 3-dimensional tertiary structure[31]. Proteins commonly assemble into oligomeric complexes that

exhibit diverse quaternary structure arrangements. The ability of fully folded protein units to assemble into a diverse set of structures have made them attractive for the design of self-assembling biomaterials.

1.4.1 Self-assembly of proteins

The primary level organization is the linear sequence of amino acids, called the peptide chain. The peptide chain folds into a second level of the organization through hydrogen bonding and other covalent interactions, the β-sheets, and α-helices. Folding of the peptide chains into an energetically favorable conformation constitutes the third level of organization, the 3-dimensional tertiary structure. The fourth level of organization is the self-assembly of fully folded protein units through non-covalent interactions to give rise to the quaternary structure[32]. Figure 1.3 shows the different levels of organization of proteins.

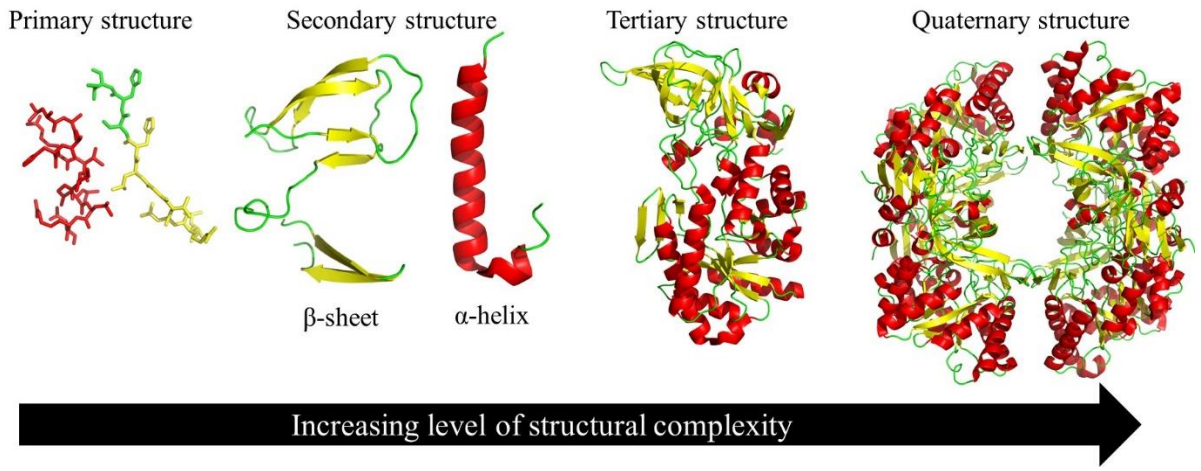


Figure 1.3: *Schematic showing the different levels of organization of protein (PDB: 2O0T).*

Most of the characterized proteins reported in literature organize themselves in quaternary structures either as homo-oligomers (proteins made of identical polypeptide chains) or hetero-oligomers (proteins made of different polypeptide chains). These higher order protein

structures are typically maintained through multiple weak non-covalent interactions, sometimes including covalent interactions[33]. Oligomerization of protein into the supramolecular quaternary complex exhibits various advantages such as higher order complexity, allosteric regulation, functional control, and stability. Such complex protein assemblies can result from either permanent or transient interactions[34]. The subtle equilibria that allow proteins to achieve diverse supramolecular architecture rely on their flexibility to undergo large conformational changes in solution[33]. Visualization of such protein assemblies *in vitro* has been possible through a range of low- and high-resolution techniques such as X-ray crystallography, nuclear magnetic resonance (NMR) spectroscopy and electron microscopy (EM)[35,36].

More recently, cryo-EM has made it possible to observe such native assemblies in solution at angstrom resolution[37]. The different types of naturally occurring architectures of proteins characterized in the context of self-assembly include linear assemblies, rings, tubes, catenanes, knots, cages, and vaults[38]. Important examples of linear protein assemblies are the collagen fibers that make up mammalian connective tissue[39] and keratin fibers that make up the nail, hooves, horns, feathers, hair and wool fibers[40].

Ring shaped protein complexes are another important assembly of proteins and certain ring shaped complexes show DNA/RNA binding and are connected with DNA and RNA manipulations. Important examples include the DNA helicases[41] and the RNA-binding attenuation protein (TRAP) involved in the termination of transcription of *trp* operon that forms an undecameric ring structure with a pore diameter of 25 Å[42]. An example of homo-oligomeric protein that self-assembles into different oligomer complexes is 2-Cys peroxiredoxins (Prxs) that protect the cells against oxidative damage[43].

Prx protein is a stable homodimer that self-assembles into ring-shaped oligomers of different sizes including, octamer, decamer and dodecamer conformations, clusters, cages, and tubes[44]. Cylindrical assembly of proteins consisting of stacked rings have been described in the case of chaperonins[45] and peroxiredoxin[46,47]. Tubular protein assemblies are another class of cylindrical protein self-assembly that were identified in the context of host infection and storage of genetic material in viruses[48], and the transmembrane channels of *S. aureus* heptameric α-hemolysin protein complex that is 100 Å in length and has a pore diameter of 14-46 Å[49]. Other interesting natural topologies encountered include the knotted proteins[50], vaults formed by ribonucleoproteins[51] and bacterial nano-compartments formed by encapsulin[52].

Many proteins that form oligomer complexes and serve different functions within living systems are rich in a specific type of secondary structure element. For example, the natural cross-β rich conformation of peptides and protein hormones stored in the secretory granules of the endocrine system[53] and in the case of amyloid fibrils as a consequence of protein misfolding[54].

1.4.2 Self-assembly through misfolding

The folding and misfolding of protein is a thermodynamically driven process in which the native state of the protein corresponds to the most thermodynamically stable conformation under physiological conditions. Proteins can achieve a variety of conformations ranging from unfolded peptide chains to partially folded intermediates and fully folded native conformation. The lowest energy state is the local minima in the highly complex energy landscape of the protein[55] (Figure 1.4). There are numerous local minima within the protein energy landscape and a change in the physiological conditions such as pH, ionic strength, temperature, and concentration can significantly influence the local minima[56].

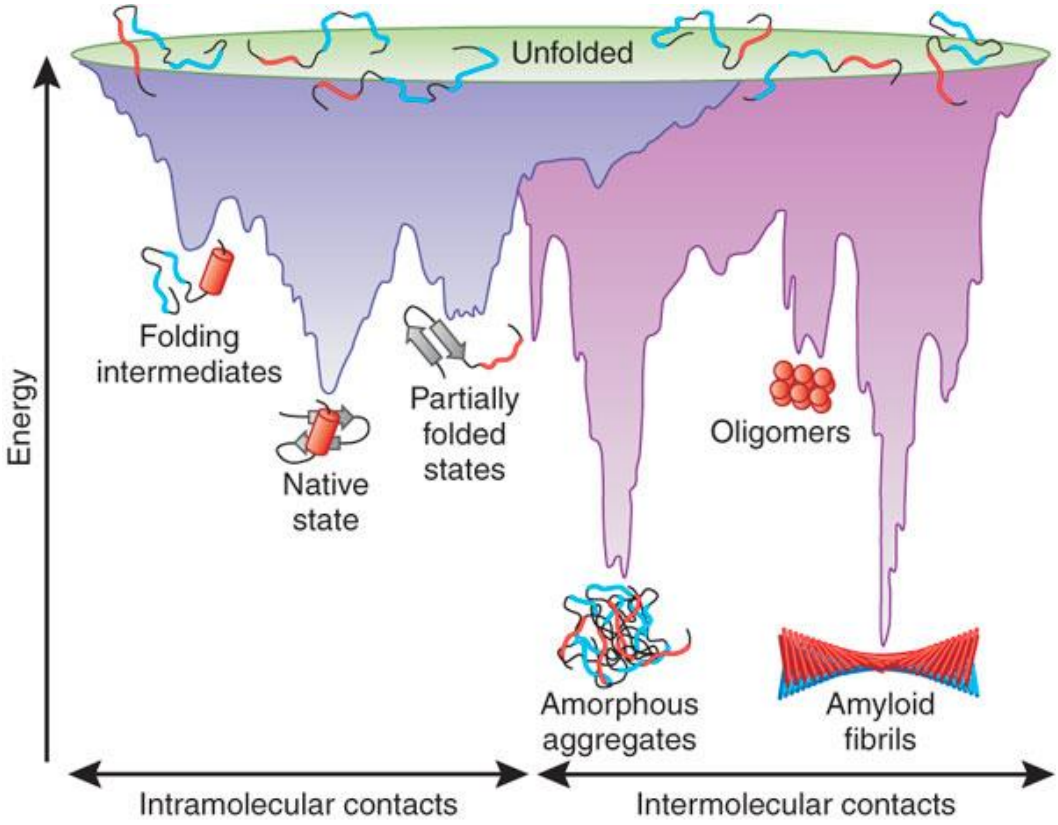


Figure 1.4: *Energy landscape of protein folding and aggregation. The purple surface shows the multitude of conformations via intramolecular contacts and the pink area shows conformations moving towards amorphous aggregates or the ordered amyloid fibrils via intermolecular contacts. Aggregate formation can occur from intermediates populated during de novo folding or by destabilization of the native state into partially folded states. Diagram taken with permission from Hartl et al. (2009)*[55].

Depending on the external environment, aggregation of protein can be disordered leading to amorphous aggregates or result in highly ordered aggregates in which the peptide or polypeptide chains are arranged in a specific fashion that is a hallmark of aggregated protein assemblies[54] (Figure 1.5).

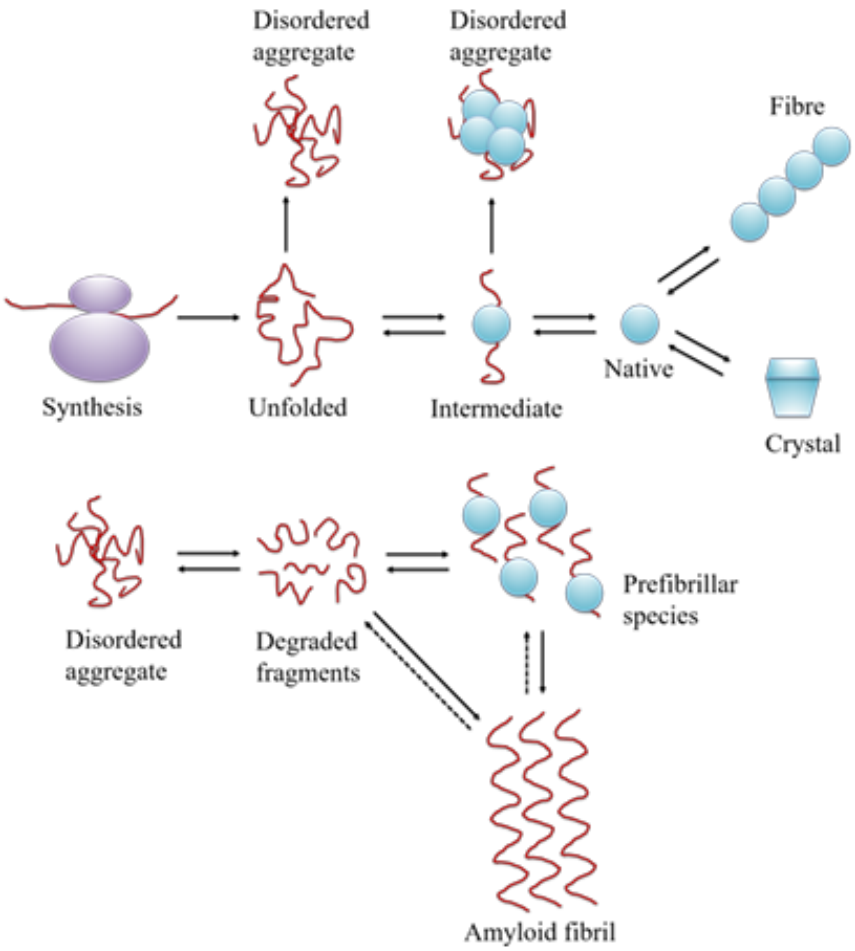


Figure 1.5: *Protein self-assembly: An overview of the different types of structures formed by peptide/polypeptide chains*[54]*. Reproduced with permission from Chiti et al. (2006)*[54]*.*

1.4.3 Amyloid fibrils – A generic protein assembly

Amyloid fibrils are insoluble fibrous, proteinaceous aggregate deposits, structurally dominated by an extended β-sheets[57]. Amyloid fibrils are associated with protein misfolding diseases such as Alzheimer's, Huntington's, Creutzfeldt-Jakob, and type II diabetes[54,58,59]. Though largely associated with diseases, functional amyloid assemblies are evident in nature in varied life forms[60] (Table 1.1). The diversity in the sequence of amyloid-forming proteins and the existence of non-disease associated proteins to form amyloid fibrils suggests that the amyloid state is possibly a generic protein fold of all polypeptide chains[69].

Table 1.1: *Native protein assemblies rich in β-sheets*

Protein	Role	Assembly	Reference
TMV capsid protein	Encapsulates the viral RNA and protects from enzymes.	Nanotubes	61
Bacterial curli proteins	Produce biofilms that help in surface adhesion and colony formation.	Fibrillar Networks	62
ChpE and ChpH Chaplin proteins	Produce fibrous amphipathic films that facilitate the aerial growth of hyphae for spore formation.	Fibers	63
Fungal class I hydrophobins	Fibrous networks on fungal coat help to adhere to hydrophobic surfaces to facilitating pathogenesis.	Fibers	64
Microbial gas vesicle protein GvpA and GvpC	Function as floatation organelle	Fibrillar Networks	65
Chorion protein	Makeup eggshell layers in insect eggs and serve a protective function	Fibers	66
Spidroin protein	Major component of the spider web matrix	Fibers	67
Pmel17 protein	Important for production of mature melanin	Fibrils	68

Amyloid fibril formation is posited to occur through nucleated polymerization[70]. Nucleated polymerization involves the growth of filamentous structures through development of growth nuclei (primary nucleation). Development of the growth nuclei occurs through individual nucleating events which creates a templates for the conformational induction of new monomers[70,71]. Time course studies on amyloid fibril formation using techniques such as Thioflavin-T (ThT) binding assays, light scattering, and electronic microscopy show the presence of a lag phase, exponential growth phase, and stationary phase. The lag phase observed is thought to be due to the formation of nuclei that are largely β-sheet seeds. Once formed, the growth and extension of fibrils occur rapidly through the addition of monomers or oligomers[71,72,73]. A three state equilibrium is proposed for amyloid forming peptides: the monomer, oligomer, and fibrils[74]. Similarly, with other nucleation-dependent assembly, the

equilibrium between monomer to fibrils can be changed by the addition of preformed fibrils or "seeds" which shortens the lag phase in assembly[71,72] (Figure 1.6).

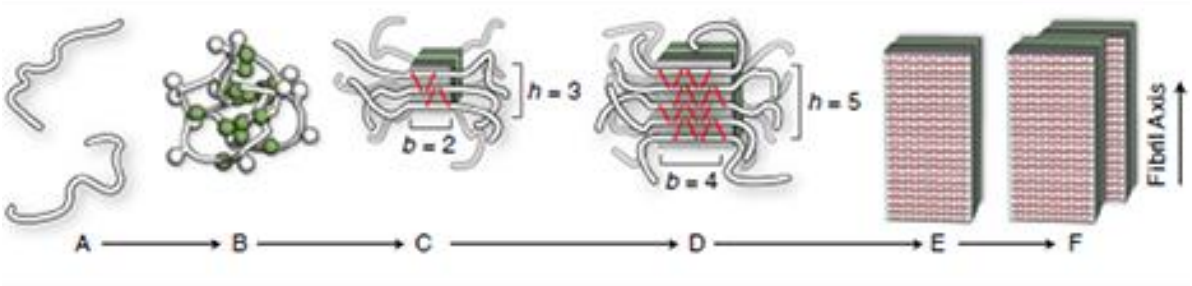


Figure 1.6: *Self-assembly process of amyloid fibril formation through nucleation: State A: Peptide monomer isolated in solution. State B: Oligomer complex of few peptide chains. State C: Nucleus of β-sheet structure. The peptide backbone runs perpendicular to the fiber axis. State D: Postcritical nucleus structure is showing more β-sheet structure. State E: The protofilament is a single long thread of β-structure consisting of a β-sandwich and two face-to-face β-sheet planes. State F; The full fibril, a bundle of protofilaments shown here to contain p = two protofilament threads[74]. Taken with permission from Schmit et al. (2011)[74].*

1.4.4 Amyloid fibril structure

Preliminary structures of amyloid fibrils came from transmission electron microscopy (TEM)[75], X-ray fiber diffraction[76], atomic force spectroscopy (AFM), and X-ray crystallography[77]. These techniques provided important information about the size, shape, unbranched morphology, presence of the characteristic repetitive cross-β architecture and polymorphism in fibrils[78,79,80,81]. Amyloid fibrils are straight, unbranched fibers with a typical diameter of 70-120 Å and several hundred nanometers in length. The cross-β architecture of amyloids is a generic feature observed for different amyloid fibrils in which the β-strands are organized perpendicular to the fiber axis[78,79]. More recently, cryo-EM coupled with 3D reconstruction provided a high-resolution model of amyloid fibrils up to 8 Å[82] (Figure 1.7).

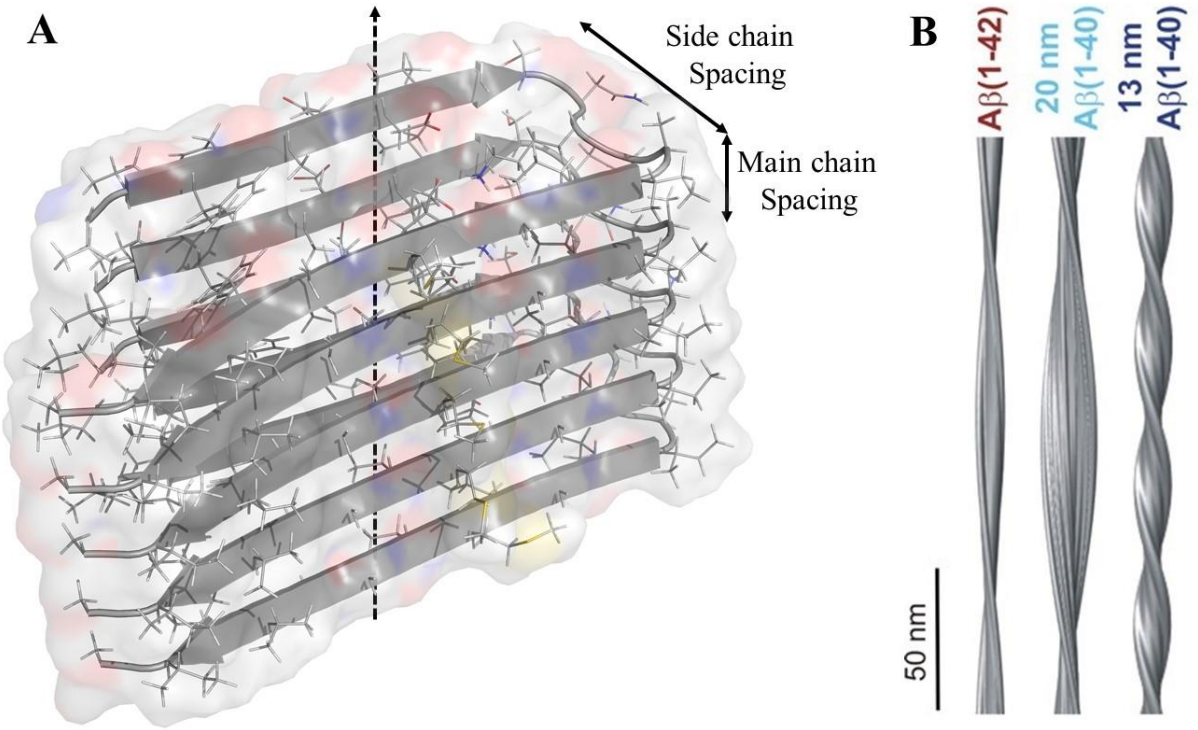


Figure 1.7: *Structure of amyloid fibrils: (A) 3D structure of amyloid fibril of $A\beta_{1\text{-}42}$ (PDB: 3BEG). (B) side views of reconstructions of Aβ (1-42) and Aβ (1-40) amyloid fibrils, taken with permission from Schmidt et al. (2009)*[82].

Solid state nuclear magnetic resonance (ssNMR) data supplemented with low-resolution structures from electron microscopy allowed for the development of molecular models of full-length amyloid fibrils[83]. Atomic structures of amyloid fibrils have been difficult to determine due to their non-crystalline and insoluble nature.

The first atomic-level structure of amyloid fibrils was published in 2005[84]. Different atomic models from X-ray crystallography of short peptide sequences that self-assemble into amyloid fibrils suggest that all amyloid fibrils share a common cross-β structure with different arrangements of β-sheets and peptide monomers within β-sheets[83,85].

One significant feature of amyloid fibrils highlighted utilizing contemporary technology is the polymorphism of amyloid fibrils. Polymorphism in amyloid fibrils is defined as the presence of different forms of fibrils resulting from substantially different arrangements of the β-strands in the fibrils[81]. A high-resolution model of $A\beta_{1-40}$ amyloid fibrils obtained using ssNMR revealed polymorphic amyloid fibrils having different folds for mature fibrils[86]. Such polymorphism had been earlier noted in amylin fibrils[87] and also consistently observed in many new peptide assemblies[88]. Polymorphism in the fibril structure is thought to arise due to different inter-sheet and intra-sheet packing of the peptide strands that gives rise to staggered arrangements with the side-chain of the terminal residues exposed. These exposed arms of the residues interact with each other to give rise to ultrastructural differences in the fibril morphology[89]. These structural constraints that are typical of amyloid fibrils make them more malleable.

Structural polymorphism is noted not only in amyloid fibrils of different proteins or peptides but diverse polymorphic amyloid fibrils are identified within same amyloid forming protein. Therefore, it is unlikely that a single model can provide an accurate view of all amyloid fibril-forming proteins and peptides.

More recently magic angle spinning (MAS) NMR spectroscopy coupled with high-resolution electron density maps from cryo-EM microscopy, X-fiber diffraction along with AFM measurements illuminated the molecular basis of stability of polymorphic amyloid fibrils formed by self-assembling peptide sequences[90]. The MAS-NMR structure fitted to the cryo-EM reconstruction revealed twisted amyloid fibrils made up of two, three or four in register β-sheets, laterally associated arranged in the extended fibril structure (Figure 1.8).

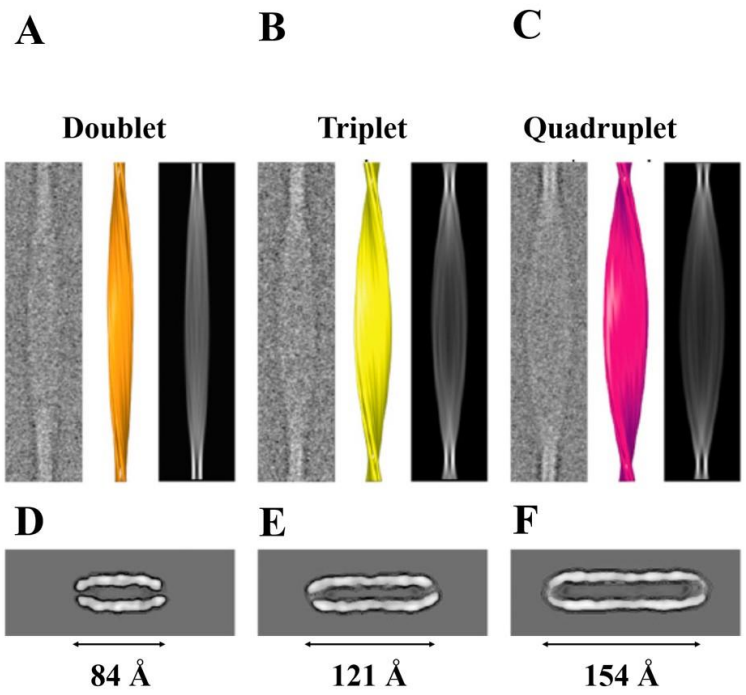


Figure 1.8: *Representative cryo-EM images of averaged fibrils (class averages), surface representations of reconstructions, 2D projections, and contoured density cross-sections of the three types of fibril formed by TTR (105–115). (A) Doublet class average; (B) Triplet class average; (C) Quadruplet class average on the left. 3D reconstructions are shown in the center (Doublet: orange, Triplet: yellow, and Quadruplet: purple).2D projections of the fibril reconstruction are shown in right. (D–F) Contoured density cross-sections of the doublet (D), triplet (E), and quadruplet (F) fibrils*[89]*. Taken with permission from Fitzpatrick et al. (2013)*[89]*.*

Amyloid fibril assembly shows a near perfect "design" and is estimated to have only one misalignment per 30,000 strands[91]. Amyloid fibrils show resistance to denaturants, detergents, thermal treatment and proteases[57]. These properties make amyloid fibril assemblies highly attractive candidates for use as as biomaterial in the emerging field of bionanotechnology for fabrication of nanostructures serving specific functions.

1.5 Bionanotechnology

Bionanotechnology is a branch of nanotechnology that links biology to physics and deals with the study of biological molecules and macromolecules at the nano-level[92]. The field aims to utilize biological designs and use biological or bio-inspired starting materials to build nanomaterials. This thesis focuses on peptide and protein-based nanostructures. The increasing interest in designing building blocks that can self-assemble into ordered hierarchical assemblies similar to those formed by protein and peptides in nature have suggested several design strategies.

Two main approaches define the design strategy for production of well-ordered nanostructures *in vitro* (Figure 1.9). The top-down approach is the processing of a larger piece of material using fining tools to create a small structure that can be used for an application. In contrast, the bottom-up approach uses smaller components of atomic/molecular dimension that self-assemble naturally or through the application of an external driving force giving rise to large and more ordered system[93]. The process of simple building blocks known as tectons coming together to form nanoscale structures is the basis of the bottom-up approach. It relies on the tendency of the building blocks to interact with each other in a coordinated way to form large and complex structures. The natural self-assembly of peptides and proteins observed in biological systems[95,96] provides a good starting point to explore their potential *in vitro* for this purpose. It is generally aggreed that “bottom up” approach is relatively cheap and energy efficient for the mass production of artificial constructs[97,98].

Nucleic acids, lipids, carbohydrates, and proteins have all been explored as building blocks to create non-biological architectures.

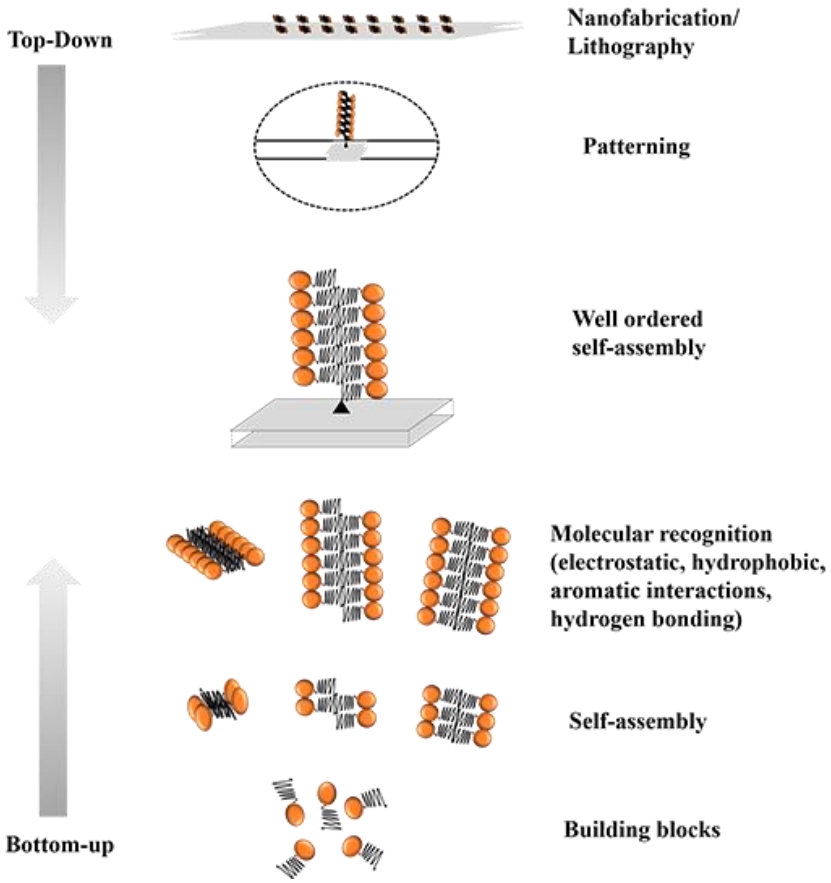


Figure 1.9: *A comparison of the "top-down" and "bottom-up" approaches applied to produce self-assembling biomaterials. The top-down approach is the patterning of assemblies on a surface by nanofabrication/lithography. The basis of the bottom-up approach is the self-assembly of simple building blocks through molecular recognition and self-assembly. Reproduced with permission from Gazit et al. (2008)*[94].

Nucleic acids offer a simple mechanism of interaction and assembly, and possess the spatial periodicity like the DNA double helix to build supramolecular structures with functionality[99]. DNA origami has been a useful tool to produce tailored building blocks with predictable self-assembly. DNA based nanostructures are attractive as scaffolds to conjugate other biomolecules that can be spatially displayed with high precision to produce rigid scaffolds with functionality or switchable architectures. For example, DNA based polyhedron structures were successfully used as delivery vehicles to ferry encapsulated molecular payload to cell *in vivo*.

DNAzymes or RNAzymes that show change in fluorescence upon activation in the presence of metal ion have been successful as tools for ionsensing. DNA aptamers that have found utility in diagnostics and biosensing applications, such as the DNA aptamer linked to antibody for reporter-linked aptamer assay to detect vascular endothelial growth factor (VEGF) and the adenosine binding aptamer to screen for inhibitors of adenosine deaminase (ADA)[99,100].. Lipid based nanostructures have been successful in nanomedicine as delivery vehicles, for example the cationic multilamellar vesicle (MLV) liposomes containing potent expression promoter-enhancer and transcriptional activator high mobility chromatic protein as gene therapy vector in metastatic breast cancer mice models[101]. DNA and lipid based nanostructures are promising for different applications in bionanotechnology however; they lack the advantage of functionality and higher order complexity of 3D structures offered by proteins. Significant success has been achieved with bioinspired and synthetic protein chips (peptide sequences) with a propensity to self-assemble into a peptide hydrogel biomaterial[102], nanofibrils or nanowires for fabrication of biosensors[103,104], tissue engineering, and 3D cell culture[105].

1.6 Protein as building blocks of self-assembly

Proteins have long been known as versatile biological building blocks. Their ability to form a range of different 3-dimensional architectures is unquestioned as nature is filled with a large number of such examples (section 1.4.1). It is relatively cheap and easy to produce such protein building blocks in the laboratory using synthetic or recombinant methods. An algorithm based approach has been successfully utilized to create a library of nanomaterials with unique configurations, and this has also allowing a large number of energetically favorable conformations to be tested[106].

Using special design programs like the Rosetta and extending it for modeling protein-protein interface interactions has allowed for the production of protein tectons with different symmetries combined in specific geometries[107,108]. Proteins usually assemble through a modular approach with protein units self-assembling into well-ordered oligomer complexes through both permanent and transient interactions[34,109]. Taking advantage of gene mutation tools and the intrinsic tolerance of many proteins to mutation non-biological assemblies of natural proteins have been created. For example, the engineered mutant of heat stable bacterial TRAP protein that self-assembles into a 12-mer protein ring with a pore diameter of 2 nm[110]. The ring structure has been further modified through the incorporation of disulfide and hydrophobic interactions to create extended nanotubes[111].

With the advent of new methods in protein engineering, it is now possible to engineer proteins incorporated with specific metal binding sites. For example, peroxiredoxin containing a peptide tag with six or more consecutive histidine residues (6*His tag) at the N-terminal end was used to sort Ni^{2+} functionalized gold nanoparticles into 1D clusters[112]. In a more recent effort to create designer material, the human SMAC protein that self-assembles as a V-shaped oligomer has been utilized to produce stable 1D zigzag nanowires and 2D wavy nanostructures through the incorporation of bis-histidine motifs that direct stable metal mediated interaction with Zn^{2+} ions[113].

Increasing interest has been shown for the development of biopolymers from renewable sources and more recently from nonfood sources. Amongst them, keratins that make up the wool fibers, feathers and horns, fibroin fibers and silk fibers have proven to be useful for production of nanofibers with the eventuality of nanofabrication. Electrospinning has proven to be highly useful for creating nanofibers with controlled fiber diameter[114,115]. Keratin based

nanofibers[116], films[117] have been used to filter heavy metal[118] or volatile organics and scaffolds in tissue engineering[119]. Gerrard *et al.* presents an excellent overview of different approaches to building protein-based nanostructures and their applications in applied biology in varied ways[97].

1.7 Peptides as building blocks of self-assembly

Peptide sequences have been designed from different unrelated native protein sequences with motifs that exhibit self-assembling properties and associate to form well-ordered hierarchical structures. Much work has been carried out with amyloid forming sequences[120,121], peptide amphiphiles[7], peptide-polymer block copolymers[122] and coiled-coil motifs[123]. Peptides can be customized due to their small size, short length and ease of large scale production.

Increasing efforts in the direction of *de novo* peptide design have introduced a library of peptide tectons with designs inspired from different biological systems, proteins motifs, and synthetic modeling[124]. The current library of peptide sequences known to self-assemble belong to two classes: one that shows a propensity to form extended β-sheets and the second that form α-helical assemblies.

1.7.1 Self-assembly of β-sheet forming peptides

The β-sheet forming peptide sequences first characterized in the context of human pathologies were found to from extended fibrils by organizing themselves into extended β-sheets laminated onto each other[54,57]. Peptide sequences that self-assemble into extended β-sheet based nanostructures include nanofibrils, nanowires, nanotubes, nanovesicles and nanosheets. They exhibit the defining characteristics described for amyloid fibrils (section 1.4.3 and 1.4.4).

High-resolution models developed from X-ray crystallography of short peptide sequences show the peptide β-strands organized into ladders of β-sheets stabilized by hydrogen bonding between neighboring amide backbone groups and individual complementary interacting residues. The in-register alignment of matching residues generates very tight packing and also maximizes the favorable hydrophobic and van der Waals side-chain contacts along the assembly axis[84,85]. The hierarchical structure called the "steric zipper" has two extended β-sheet laminated onto each other with a wet (hydrophilic) and a dry (hydrophobic) interface, with polar residues of the wet interface hydrated with water molecules and that of the dry interface tightly interdigitated complementary to each other[84].

The different organizations of the peptide sequences is classified into eight different types of steric zipper classes based on the arrangement of the β-strands in the sheet and the orientation of the β-sheets in the extended amyloid fibrils (Figure 1.10).

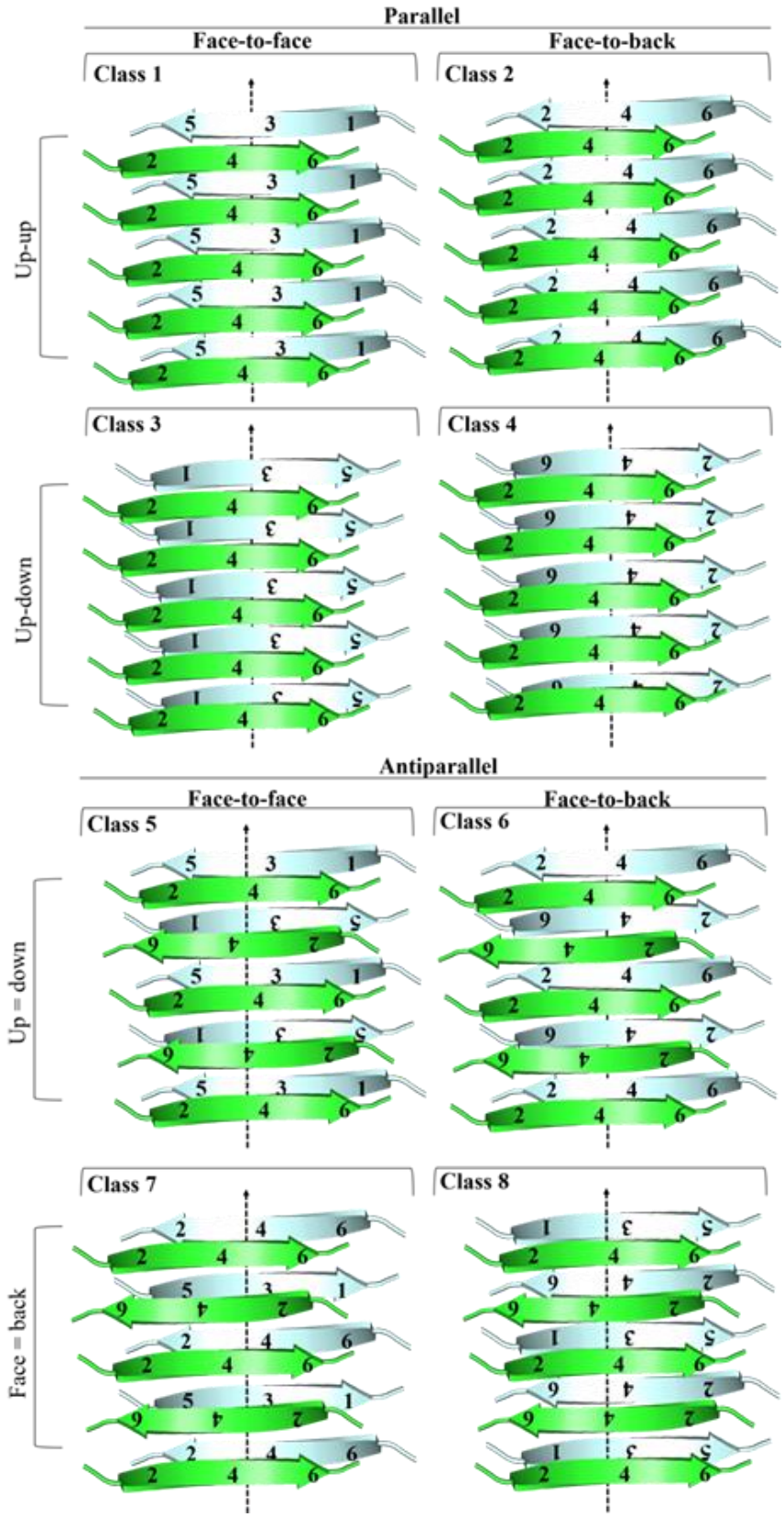


Figure 1.10: *Cartoon model representation of the eight different classes of steric zippers. Reproduced with permission from Sawaya et al. (2007)*[125].

1.7.2 Defining characteristics of β-sheet peptide assembly

General methods used for recognizing amyloid or amyloid-like assembly includes binding of fluorescent dyes such as Congo Red (CR) and ThT. Defining characteristics of the self-assembled extended peptide nanostructures can be resolved by AFM and TEM, the β-sheet secondary structure is measured by circular dichroism (CD) spectroscopy and Fourier transform infrared (IR) spectroscopy[126,127], and the organization of β-sheets within the fibrils done by X-ray diffraction and scattering techniques[128].

X-ray fiber diffraction is a standard method to characterize amyloid fibrils, which give characteristic diffraction patterns that correspond to the specific alignment distances of the peptide β-strands in the cross-β structure. The specific reflection pattern at 4.8 Å in the meridional region corresponds to the hydrogen bonding distance between the side chains in one β-sheet. The diffuse reflection at 10-11 Å in the equatorial region corresponds to the distance between the β-sheets[129,130] (Figure 1.11). Successive studies on aggregating peptides designed from amyloidogenic motifs within proteins have elucidated a third reflection observed at 9.6 Å specific to the repeat distance between the anti-parallel orientations of the β-strands[131].

Small angle X-ray scattering (SAXS) and wide angle X-ray scattering (WAXS) have proven to be useful tools that can provide molecular level information on extended protein/peptide assemblies under hydrated and physiological conditions within relevant millisecond time scale. SAXS measures the elastic scaterring of X-rays by a sample with inhomogeneities in the nm range and is generally recorded at small angles (typically 0 to 10°). WAXS is similar to SAXS, however the distance of the sample to the detector in shorter and thus the diffraction maxima is observed at larger angles[128].

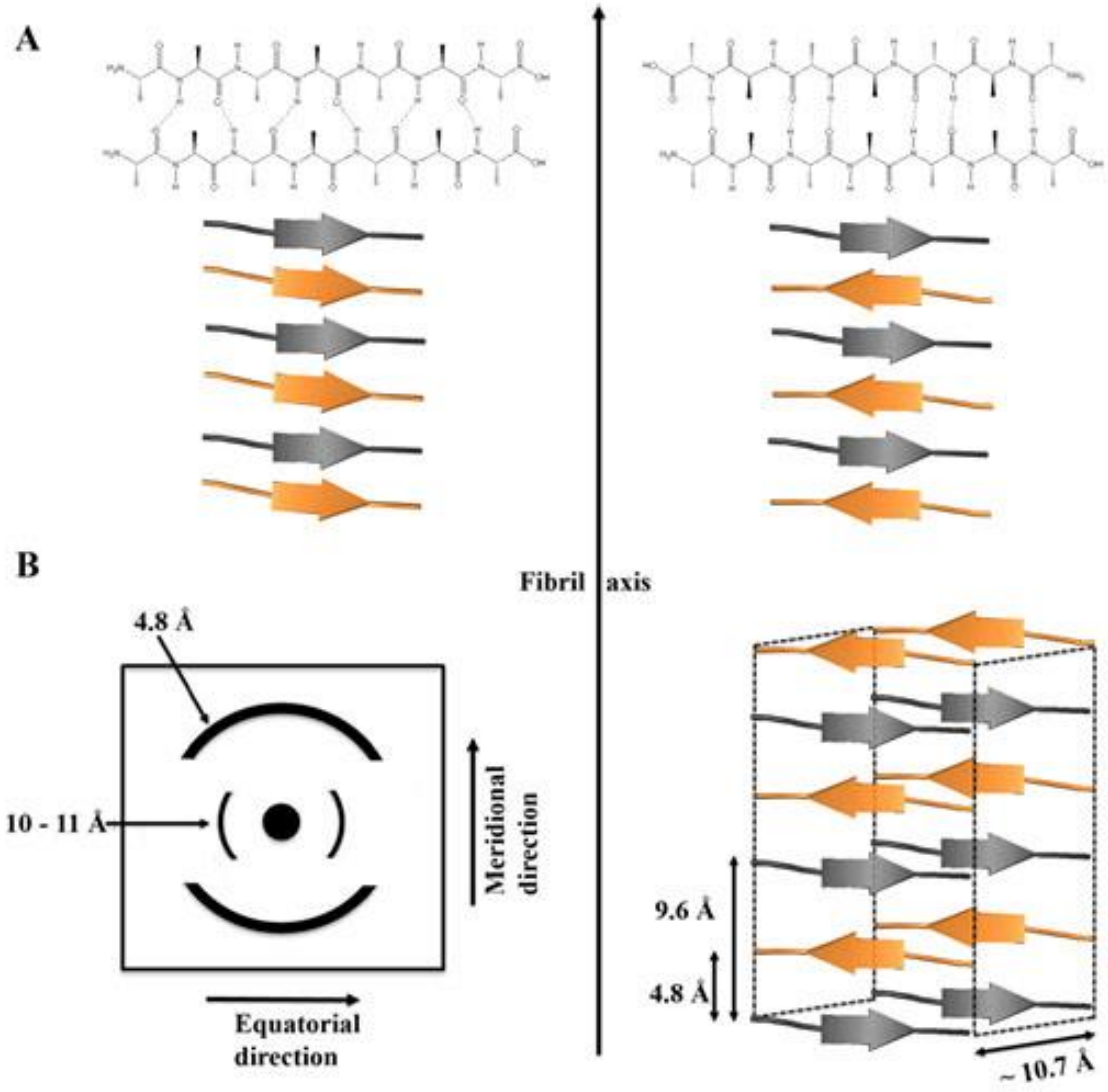


Figure 1.11: *Structure of β-sheets and X-ray diffraction of amyloid fibril: (A). Schematic of the assembly of β-strand motifs into ordered parallel and antiparallel β-sheets. The chemical structure shows the backbone hydrogen bonding*[132]*. (B) X-ray fiber diffraction model from amyloid fibrils showing the major reflections. The cartoon model shows the anti-parallel cross-β architecture that is the hallmark of amyloid fibril structure. Reproduced with permission from Serpell et al. (2000)*[133]*.*

The techniques have proven to be really useful in elucidating that the strongest repetitive structure in the amyloid fibril architecture is the set of β-sheets aligned parallel to the fibril axis and the β-strands perpendicular to the fibril axis[133]. Studies on self-assembling β-sheet forming

peptides showed that X-ray scattering at small and wide angles can be used to model the arrangement of β-sheets in the large embedded architectures like peptide nanotubes[134].

Congo red (CR) is an amphiphilic dye and are widely used for staining and detection of amyloid fibril assembly. CR shows absorption maxima at 490 nm in aqueous solution leading to the characteristic red color solution. CR molecules upon binding to β-sheet rich amyloid fibrils become torsionally restricted (long axis of the CR molecules oriented parallel to the fibril axis). This induces a characteristic increase in the absorption as well as a red shift in the absorption maximum from 490 to 540 nm. Furthermore, this type of binding results in green birefringence when observed under cross-polarized light[135].

ThT is a benzothiazole dye that exhibits enhanced fluorescence intensity due to the shift of the excitation maximum from 385 nm to ~ 450 nm and emission maximum from 445 nm to ~ 482 nm upon binding to extended β-sheet assemblies like amyloid fibrils[136]. ThT binds to diverse fibrils despite their distinct amino acid sequences which suggest that it recognizes a structural feature that is generic to amyloid fibrils. Recent theory suggests that ThT binds to the β-sheet surface along cavities running parallel to the fibril axis, between the protofilaments[136,137]. Binding modes of ThT to amyloid fibrils studied by NMR show that ThT binds orthogonal to the fibrils and complexes differently to fibrils of different proteins[138]. Other fluorescent dyes such as amino derivatives of benzanthrone, squaraine dye SQ-1, and polymethine dye V2 are currently being investigated as alternate fluorescent probes due to their property to associate with amyloid assemblies with greater sensitivity and relatively lower background fluorescence[139].

TEM is an effective technique to visualize and distinguish nanostructures of different morphologies from disordered peptide and protein aggregates. In TEM, a beam of electrons is produced using a hot filament and focused on the sample. Under TEM, amyloid-like fibrillar assemblies usually appear as extended nanostructures that are unbranched, twisted in appearance, may be intertwined, with dimensions of 6-20 nm in width and several microns in lengths[140].

Aside from TEM, scanning electron microscopy (SEM) and Atomic force microscopy (AFM) are widely used to image amyloid fibrils. In SEM, the surface of the sample is scanned by focusing the electron beam on the sample using deflection coils and focusing lenses. SEM has been notably helpful for detection of small and thin amyloid fibrils of diameters as small as 3.5 nm, as the apparent width of amyloid fibrils observed by SEM is much broader than the original value. In AFM on the other hand a cantilever tip scans the surface contour of the sample adsorbed onto atomically flat smooth surfaces (typically mica), and provides high resolution topography (≤ 1 nm) of the sample. AFM provides an added advantage to TEM and SEM to visualize amyloid fibril assembly, as it allows for continuous monitoring of the sample. This has been really useful for monitoring and imaging the growth of oligomer and fibrils in solution[140].

For the experiments in this thesis, the fluorescent dye ThT along with CD spectroscopy, SAXS and FTIR was used to detect and estimate the extended β-sheet based amyloid-like fibrillar assemblies in solution. TEM was used to visualize both peptide and protein self-assembled nanostructures[126,33]. Also, EM micrographs were used for measurement of the nanostructures using the software ImageJ[141].

1.7.3 Current sources of self-assembling β-sheet type peptides

The current library of peptides self-assembling into extended β-sheet based structures includes sequences designed from functional amyloids, amyloid-forming polypeptides and proteins, structural proteins, antibiotics, natural biological assemblies and de novo designs made from permutation of these sequences[124].The next section presents an overview of the different peptide designs that self-assemble into extended β-sheet based structures.

1.7.3.1 Functional amyloid inspired peptides

Some peptides that function as hormones are thought to be stored in secretory granules, with an amyloid-like cross-β-sheet rich conformation[53]. Somatostatin-14, a small cyclic neuropeptide that regulates the growth hormone glucagon, insulin and gastric secretions has a high propensity to form antiparallel β-sheet structures in water[142]. Valery & co-workers designed the peptide Lanreotide, an analog of somatostatin hormone; that self-assembles into liquid crystalline nanotubes in aqueous solution[143]. The peptide formulation is currently marketed as a treatment for acromegaly[144].

1.7.3.2 Amyloid-inspired peptides

Exploiting the natural amyloidogenic sequences, a library of peptide sequences have been designed that self-assemble in solution into different nanoscale assemblies. A variety of sequences able to form amyloid-like fibrils raised the idea of a generic feature coming from the intrinsic properties of peptide chains under appropriate conditions. Short peptide sequences comprised of 4-6 amino acid residues self-assemble into typical amyloid fibrils formed by a full-length polypeptide chain. Examples include, the hexapeptide “NFGAIL” derived from the islet amyloid polypeptide (IAPP)[145], the pentapeptide “DFNKF” derived from the human

calcitonin[146] and hexapeptide "NFGSVQ" derived from human medin[147], that self-assembles into fibrils with ultrastructure and biophysical properties of amyloid structures.

An important observation in many amyloid inspired self-assembling peptide sequences is the presence of phenylalanine residue. It is considered to be an important residue involved in amyloid fibril formation through aromatic stacking. Studies on variant peptide sequences designed with the phenylalanine residue replaced with aliphatic amino acids exhibited substantially reduced fibril-forming ability[148]. Peptide sequences comprised of one or more phenylalanine residues self-assembles into extended nanotubes, nanofibrils, and nanoribbons that are responsive to acidic pH, ionic strength, temperature and divalent metal ions[149,150,151]. Diphenylalanine nanowires have been successfully fabricated on surfaces for the design of biosensing devices[152].

1.7.3.3 Cyclic peptides

Cyclic peptides were one of the earliest set of synthetic peptides designed from natural antimicrobial peptides produced in microorganisms, plants, and animals, as part of the innate immune response[153]. The rationally designed peptide analog of gramicidin A, cyclo-[-(L-Gln-D-Ala-L-Glu-D-Ala)$_n$-] adopts a flat ring conformation which stacks through backbone hydrogen bonds to form extended cylindrical structures with a pore diameter of ~1 nm[154]. Variations in the design strategy have given rise to peptide sequences that have been utilized to produce transmembrane channel mimics[155] similar to the natural transport protein channels and cyclic D, L-α-glycopeptides with antimicrobial activity[156].

1.7.3.4 Amphiphilic peptides

Amphiphilic peptides contain hydrophobic and hydrophilic residues arranged in a two-faced design. Peptide sequences Ac-A_6D, Ac-V_6D, Ac-V_6D_2, Ac- L_6D_2 and Ac-A_6K-NH_2 reportedly self-assembled into organized nanostructures including micelles, vesicles, and tubes of nanoscale dimension with dense β-sheet networks[157,158]. The overall design of amphiphilic peptides is similar to native phospholipids in the cell membrane except they have not only hydrophobic tail packing but also intermolecular hydrogen bonds along the backbone[159]. **Peptide amphiphiles** (PA) are short peptide sequences covalently joined to an alkyl chain. PA sequences designed with hydrophobic sequences to promote β-sheet formation yielded well-ordered cylindrical nanostructures[7]. Variation of this design through the introduction of charged residues in the peptide segment allowed linking of specific chemical moieties and biologically active sequences[160,161]. Another important design is the drug amphiphiles (DA) which self-assembles into discrete extended nanofibers/nanotubes with applications as independent drug delivery vehicles[162].

1.7.3.5 Ionic self-complementary peptides

Ionic-complementary peptides have alternating arrangement of negatively and positively charged residues. To have ionic complementarity, the distribution of the charges follow specific patterns, such as +-+-+-+ or - -++- -++ or - - - - ++++. A prominent example is the EAK16 peptide inspired from zoutin protein, a putative Z-DNA binding protein in yeast that self-assembles into highly pH tolerant thermostable nanofibrils. Peptide sequences with the alanine residue changed to hydrophobic residues self-assembled into interwoven nanofibers and nanopores with a diameter between 50-200 nm[163]. Zhang & co-workers utilized this design strategy to create molecular switches with the ability to change their molecular structure under different conditions of temperature and pH[164].

1.7.3.6 Multi-domain peptides

A set of interesting peptide sequences has been designed by combining domains with different properties such as the ABA block motif with the B block comprised of alternating hydrophilic and hydrophobic amino acid residues surrounded by the A block made of charged residues. Examples include the $K_2(QL)_6K_2$, $K_2(SL)_6K_2$, $E(QL)_6E$ and $E(SL)_6E$ sequences that self-assemble into well-defined nanofibers[165]. Table 1.2 presents some important biologically inspired peptide sequences that self-assemble into extended β-sheet based nanostructures.

Table 1.2: *Examples of some important self-assembling peptides with β-sheet morphology*

Biological model system	Peptide sequence	Supramolecular assemblies	Reference
Amyloid-inspired peptides			
	$(FF)_n$	Nanotubes, nanowires	166
Aβ-amyloid peptide	KLVFFE	Amyloid Fibrils	149
	KFFEAAAKKFFE	Amyloid fibrils	167
PrP	GNNNQQNY	Amyloid Fibrils	167
Cyclic peptides			
Gramicidin A	Cyclo-(L-Q-D-A-L-Q-D-A)$_2$	Nanotubes	168
Somatostatin-14 (hormone)	Lanreotide	Nanotubes	143
Amphiphilic peptides			
Phospholipid like structure	AAAAAAD	Nanotubes	7
	GGGGGGGGGGDD	Nanovesicles	157
	VVVVVVDD	Nanotubes	159
Zoutin (yeast protein)	RADARADA	Nanofibers	159
	LELELKLKLELELKLK	Nanovesicles	169

The advantage of customizable length, permutation and combination of residues, and sequence repeats confers a range of different properties like charge, hydrophobicity and aromaticity that are highly appealing for their application and utility for development of functional bionanomaterials such as diphenylalanine peptide nanowires to template metal ions and peptide amphiphile conjugated to hydrophobic drug candidates for targeted delivery in cancer therapy.

1.7.4 Self-assembly of helical coiled coil forming peptides

The α-helix is the major secondary structure element in proteins[170]. The clockwise oriented loosely coiled spring-like structure is held stably together by a network of hydrogen bonds across the coil, with a hydrophobic core. The α-helical coiled coil represents the largest assembled motifs in proteins, made up of two, three or four right-handed amphipathic α-helices wrapped around in a left-handed super-coiled fashion with a continuous in-contact hydrophobic surface[171,172]. The distinct feature of the coiled-coil assembly is the heptad repeat sequence denoted by $(abcdefg)_n$, where positions "a" and "d" is occupied by hydrophobic amino acid residues forming the hydrophobic core and other positions occupied by hydrophilic or charged residues. The overall structure of the α-helical coiled-coil is stabilized by hydrophobic interactions along with electrostatic interactions and salt bridges arising from charged amino acid residues placed in position e and g in the heptad repeat of the neighboring helices[172] (Figure 1.12).

Another hierarchical structure formed by the coiling of α-helices are the helical bundles. Helical coiled-coil bundles are hierarchical organizations of coiled coils, wherein two or more coiled-coils are packed with a sterically close-packed hydrophobic core[173].

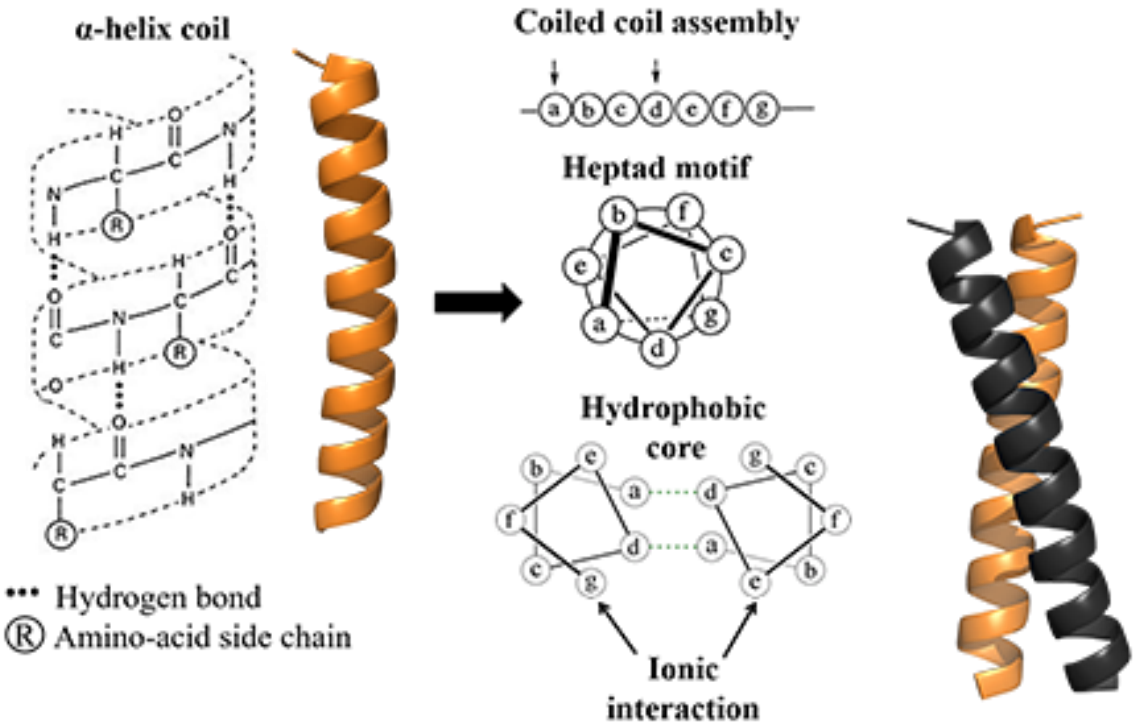


Figure 1.12: *Schematic of the α-helix coiled coil: Showing the organization of the heptad motif and the assembly of the helix coil into coiled coil. Reproduced with permission from Kohn et al. (1997)*[172].

1.7.5 Current sources of self-assembling helical type peptides

Tirell & co-workers were the first to exploit the coiled-coil assembly by designing polyelectrolyte blocks rich in alanine, glycine, proline and glutamine, flanked by dimeric coiled-coil sequences that self-assembled into pH-responsive hydrogels[174]. Kojima *et al.* introduced the first single component α-helix based self-assembling peptide in 1997. The α3-peptide with the sequence "LETLAKA" self-assembled into extended fibrils that were 5-10 nm thick[175].

Kajava & co-workers utilized this design strategy to produce peptide sequences with heptad repeats in multiples of seven, which resulted in staggered ended coiled coils that self-assembled into fibers[176]. Conticello *et al.* further explored this design to produce versatile self-assembling

fiber (SAF) type peptide systems based on the single 42-residue leucine zipper-based helix motif[177]. Woolfson & co-workers designed many peptide sequences termed magic wand peptides that self-assemble into fibrous materials[178] and added to the increasing set of SAF peptide sequences.

Kopecek *et al*. further modified the design strategy by introducing a His-tag fused to the coiled coil peptide sequence that was metal coordinated linked to a synthetic polymer. This peptide-polymer complex self-assembled into fibrous biomaterial[179]. Another noteworthy design is the GCN4 peptide inspired from the yeast transcription factor GCN4 that assembles into extended fibers with a tetrameric coiled-coil core triggered by concentration[180]. Other important designs inspired from natural systems that make up the library of the helical coiled-coil forming peptides are highlighted in table 1.3.

Table 1.3: *Current sources of helical coiled coil forming peptides*

Source	**Model**	**Assembly**	**References**
B-ZIP and GCN4 protein	Leucine zipper	Tetrameric coiled coil fibers	181,182
Collagen	$(GPO)_n$	Triple helix	183,184
Phospholipids	Diblock polypeptides	Hydrogels	185
Actin, tubulin, myosin, tropomyosin, and intermediate filament protein	Heptad repeat	Dimeric, trimeric, tetrameric and pentameric coiled coil fibers	186

1.8 Design source for novel self-assembling peptides

The various design strategies and sources discussed above have led to the development of peptide sequences that can be triggered to self-assemble into a range of different nanostructures. Furthermore, the knowledge gained from the understanding of the self-assembly process of different isolated peptide sequences has led to the development of sequences that can be tailored to meet the demands of specific applications including 3D cell culture, bio-mineralization, scaffolds for tissue repair, nanocircuitry, and delivery vehicles for bioactive molecules. The different sequences designed to date largely belong to the limited set of biological design sources highlighted above. Therefore, the need of new design sources for novel peptide sequences that can self-assemble in a regulated manner is thus warranted.

Proteins usually function as homo-oligomers. Extensive study of oligomeric protein systems have shown oligomerization to be dependent on the concentration(s), the dissociation constant(s) (K_d)[187] and change in the physiological environment, such as pH, ionic strength, and temperature[188]. Such subtle equilibria existing between monomer and oligomer forms is brought about by the protein-protein interfaces that appear to be evolutionarily optimized to regulate self-assembly[189,190,191]. Analysis of the protein interfaces reveals regions of close contacts that provide the necessary structural versatility for multiple complementary non-covalent interactions (section 1.2). The tendency of these interface motifs to interact with each other specifically as a homopolymer makes them attractive to explore for the design of self-assembling peptide sequences.

1.9 Significance and aims of the research

Self-assembling materials are widely explored for fabrication of nanodevices for different applications in bio-nanotechnology such as biosensing. templating minerals/metal ions and polymers for producing conductive nanotubes or nanowires[124]. Producing a building block that can self-assemble into stable nanostructures of defined morphology presents many challenges. Furthermore, sequences that can be utilized as components for bio-nanodevices such as bioactuators and biosensors require specific properties. Biologically inspired systems such as peptide sequences offer unique properties and features from the self-assembly perspective[92,192]. The different peptide sequence design strategies reported have allowed the creation of molecular biomaterials with better control over their dimensions and patterns of assembly[94]. The design sources for such sequences are however limited to the amyloidogenic motifs within proteins and other biologically inspired designs highlighted in the above sections. To design new self-assembling peptide sequences it is important to identify new design sources which requires understanding the natural self-assembly process of the biomolecular building blocks. Figure 1.13 shows the difference between the amyloidogenic assembly and the natural self-assembly in oligomer protein systems.

The different ways in which protein-protein interface motifs interact and arrange themselves in hierarchical structures provides a useful starting point to screen for sources of self-assembling peptide sequences. Protein-protein interfaces are evolutionary fine-tuned to allow self-assembly of protein units in a precise and controlled manner. Thus, protein interface sequences in complementary noncovalent close-contacts offer a key feature for the design of peptide sequences that self-assemble in isolation. Simple protein-protein interfaces of homo-oligomeric proteins that consist of a short stretch of complementary interacting sequences, were thus explored.

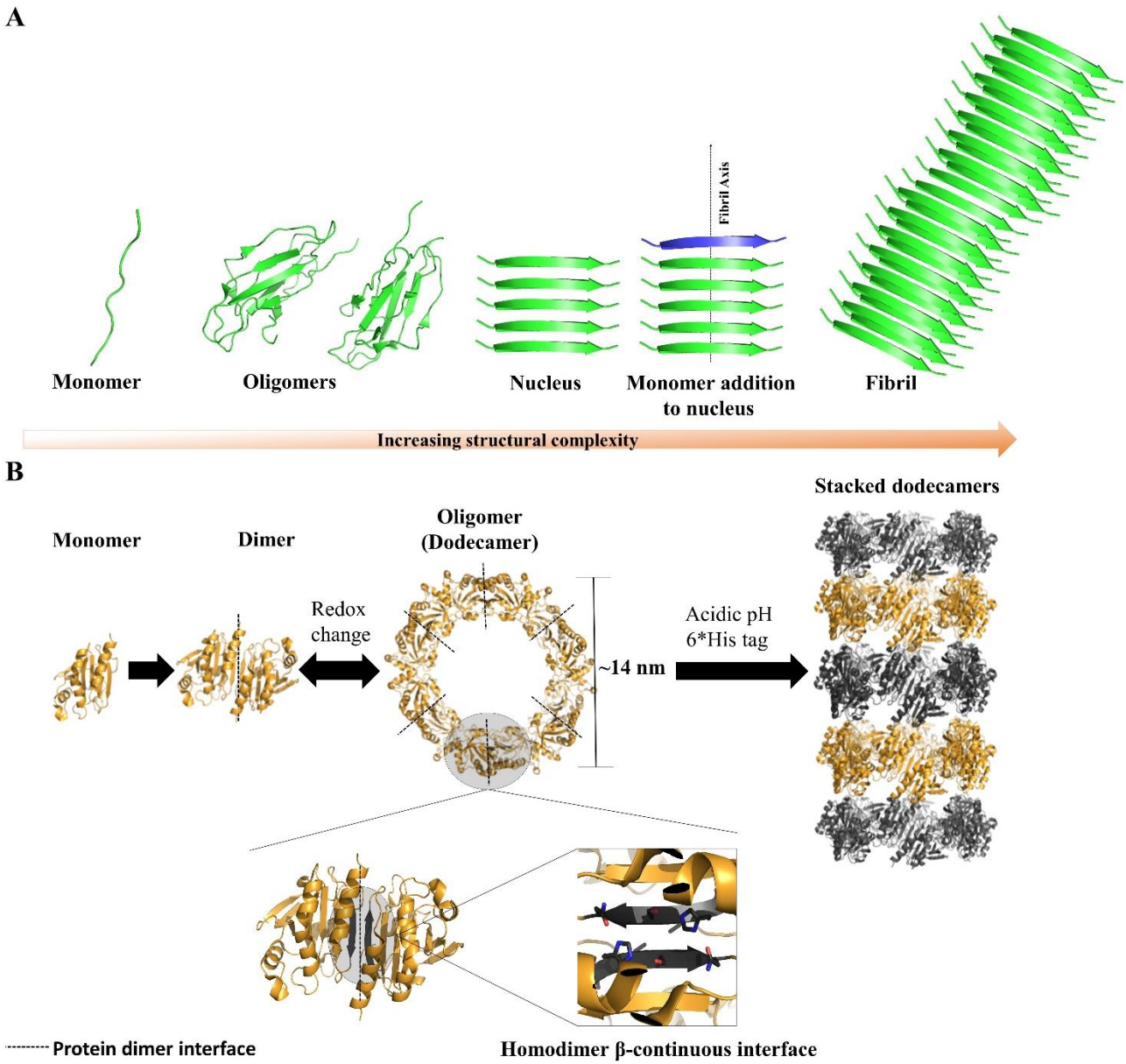


Figure 1.13: *Cartoon model showing the difference between the amyloidogenic assembly of peptide sequence inspired from transthyretin protein*[89] *(A) and the natural self-assembly of oligomer protein system bovine Prx3 (PDB: 1ZYE) (B). The simple β-continuous protein-protein interface used to design inspire peptide sequence is shown in square box.*

In this research work two different type of protein-protein interfaces were investigated for the design of self-assembling peptide sequences: the β-continuous protein interface and the helical coiled-coil interface.

The aims of the research experiments for this thesis work were:

(I) To examine protein β-continuous interface across different classes of protein homomers for the design of β-sheet forming self-assembling peptides and the characterization of the self-assembly through biophysical techniques.

(II) To investigate the role of the histidine residue in the assembly of peroxiredoxin 3 derived β-interface peptide through single residue mutation as well as to explore the parent peroxiredoxin 3 protein β-interface through site-specific mutagenesis of a histidine residue.

(III) To produce higher order assembly of human peroxiredoxin 3 (*h*Prx3) dodecamer toroids using β-sheet peptide templating and characterize the self-assembly through biophysical techniques.

(IV) To examine the coiled-coil interface of the intermediate filament forming keratin protein of *ovine* origin for the design of responsive α-helical coiled coil self-assembling peptides and characterization of the self-assembly.

1.10 Thesis overview

Chapter 1 gives an introduction to molecular self-assembly, the current sources self-assembling peptide sequences and techniques used to characterize such assembly. The chapter also introduces new source to design self-assembling peptide tectons.

Chapter 2 presents the materials utilized and experimental procedures performed throughout the research work.

Chapter 3 presents the strategy for the design of β-sheet forming self-assembling peptide sequences. The chapter discusses the self-assembly of the different peptide sequences designed and the characterization of the assemblies using different biophysical methods.

Chapter 4 details the design of variant peptide sequences and characterization of the resulting self-assembly. The chapter also discusses the different site-directed muteins of peroxiredoxin 3 protein and the induced change in the self-assembly of the oligomeric state.

Chapter 5 presents a strategy to generate higher order assembly of peroxiredoxin 3 toroid shaped oligomer into microscale assemblies and also discusses methods to control the hierarchical organization of the protein toroids into hierarchical assemblies.

Chapter 6 outlines the design strategy and introduces a new source for the design of α-helical coiled-coil forming peptide sequences. The chapter also discusses methods to achieve control over the hierarchical organization of coiled coils into extended nanostructures.

Chapter 7 discusses and summarizes the major findings in this research work and directions for future work to be pursued from this body of work.

1.9 References

1. Whitesides, G. M. & Grzybowski, B. (2002). Self-assembly at all scales. *Science* **295**, 2418-2421.
2. Armstrong, C. T., Boyle, A. L., Bromley, E. H. C., Mahmoud, Z. N., Smith, L., Thomson, A. R. & Woolfson, D. N. (2009). Rational design of peptide-based building blocks for nanoscience and synthetic biology. *Faraday Discussions* **143**, 305-317.
3. Rajagopal, K. & Schneider, J. P. (2004). Self-assembling peptides and proteins for nanotechnological applications. *Current Opinion in Structural Biology* **14**, 480-486.
4. Whitesides, G. M., Mathias, J. P. & Seto, C. T. (1991). Molecular self-assembly and nanochemistry: a chemical strategy for the synthesis of nanostructures. *Science* **254**, 1312-1319.
5. Mendes, A. C., Baran, E. T., Reis, R. L. & Azevedo, H. S. (2013). Self-assembly in nature: using the principles of nature to create complex nanobiomaterials. *Wiley Interdisciplinary Reviews: Nanomedicine and Nanobiotechnology* **5**, 582-612.
6. Atkins, P. & de Paula, J. (2011). Physical chemistry for the life sciences, *W. H. Freeman*.
7. Cui, H., Webber, M. J. & Stupp, S. I. (2010). Self-assembly of peptide amphiphiles: from molecules to nanostructures to biomaterials. *Biopolymers* **94**, 1-18.
8. Lee, Y. S. (2008). Self-Assembly and Nanotechnology: a force balance approach. *John Wiley & sons*.
9. Garrett, R. H. & Grisham, C. M. (2012). Biochemistry. *Cengage Learning*.
10. Yu, T., Lee, O. S. & Schatz, G. C. (2013). Steered molecular dynamics studies of the potential of mean force for peptide amphiphile self-assembly into cylindrical nanofibers. *The Journal of Physical Chemistry A* **117**, 7453-7460.
11. McGaughey, G. B., Gagne, M. & Rappe, A. K. (1998). pi-Stacking interactions. Alive and well in proteins. *The Journal of Biological Chemistry* **273**, 15458-15463.
12. Zayed, J. M., Nouvel, N., Rauwald, U. & Scherman, O. A. (2010). Chemical complexity-supramolecular self-assembly of synthetic and biological building blocks in water. *Chemical Society Reviews* **39**, 2806-2816.
13. Lehn, J.-M. (1988). Supramolecular Chemistry-scope and perspectives molecules, supermolecules, and molecular devices. *Angewandte Chemie International Edition in English* **27**, 89-112.

14. Lehn, J. M. (1995). Supramolecular chemistry: concepts and perspectives, *John Wiley & sons*.
15. Lehn, J.M. (1988). Supramolecular chemistry - Scope and perspectives molecules, supermolecules, molecular devices. *Journal of Inclusion Phenomena* **6**, 351-396.
16. Cram, D. J., Karbach, S., Kim, Y. H., Baczynskyj, L., Marti, K., Sampson, R. M. & Kalleymeyn, G. W. (1988). Host-guest complexation. 47. Carcerands and carcaplexes, the first closed molecular container compounds. *Journal of the American Chemical Society* **110**, 2554-2560.
17. Aoyama, Y., Tanaka, Y., Toi, H. & Ogoshi, H. (1988). Polar host-guest interaction. Binding of nonionic polar compounds with a resorcinol-aldehyde cyclooligomer as a lipophilic polar host. *Journal of the American Chemical Society* **110**, 634-635.
18. Aoyama, Y., Tanaka, Y. & Sugahara, S. (1989). Molecular recognition. 5. Molecular recognition of sugars via hydrogen-bonding interaction with a synthetic polyhydroxy macrocycle. *Journal of the American Chemical Society* **111**, 5397-5404.
19. Vicens, J., Baklouit, L. & Harrowfield, J. (2006). Calixarenes in the nanoworld, *Springer science*.
20. Evan-Salem, T., Baruch, I., Avram, L., Cohen, Y., Palmer, L. C. & Rebek, J., Jr. (2006). Resorcinarenes are hexameric capsules in solution **103**, 12296-12300.
21. Fenniri, H., Deng, B.-L., Ribbe, A. E., Hallenga, K., Jacob, J. & Thiyagarajan, P. (2002). Entropically driven self-assembly of multichannel rosette nanotubes. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 6487-6492.
22. Dubacheva, G. V., Curk, T., Mognetti, B. M., Auzély-Velty, R., Frenkel, D. & Richter, R. P. (2014). Superselective targeting using multivalent polymers. *Journal of the American Chemical Society* **136**, 1722-1725.
23. Jeon, W. S., Kim, E., Ko, Y. H., Hwang, I., Lee, J. W., Kim, S.-Y., Kim, H.-J. & Kim, K. (2005). Molecular loop lock: A redox-driven molecular machine based on a host-stabilized charge-transfer complex. *Angewandte Chemie International Edition in English* **117**, 89-93.
24. Rivera-Sánchez, M. d. C., Andújar-de-Sanctis, I., García-Arriaga, M., Gubala, V., Hobley, G. & Rivera, J. M. (2009). Walking a supramolecular tightrope: a self-assembled dodecamer from an 8-aryl-2′-deoxyguanosine derivative. *Journal of the American Chemical Society* **131**, 10403-10405.

25. Forster, S. & Plantenberg, T. (2002). From self-organizing polymers to nanohybrid and biomaterials. *Angewandte Chemie International Edition in English* **41**, 689-714.
26. Chiu, H. C., Lin, Y. W., Huang, Y. F., Chuang, C. K. & Chern, C. S. (2008). Polymer vesicles containing small vesicles within interior aqueous compartments and pH-responsive transmembrane channels. *Angewandte Chemie International Edition in English* **47**, 1875-1878.
27. Timsit, Y. & Moras, D. (1994). DNA self-fitting: the double helix directs the geometry of its supramolecular assembly. *EMBO Journal* **13**, 2737-2746.
28. Alberts B, J. A., Lewis J. (2002). The lipid bilayer. In molecular biology of the cell. *Garland Science*.
29. Israelachvili, J. N., Mitchell, D. J. & Ninham, B. W. (1977). Theory of self-assembly of lipid bilayers and vesicles. *Biochimica et Biophysica Acta: Biomembranes* **470**, 185-201.
30. Dumitriu, S. (2004). Polysaccharides: structural diversity and functional versatility. *CRC Press*.
31. Berg J. M., Tymoczko J.L., Stryer L. (2002). Chapter 3, protein structure and function. *Biochemistry*. *W H Freeman*.
32. Petsko, G. A. & Ringe, D. (2004). Protein structure and function. *New Science Press*.
33. Marsh, J. A. & Teichmann, S. A. (2015). Structure, dynamics, assembly, and evolution of protein complexes. *Annual Review of Biochemistry* **84**, 551-575.
34. Nooren, I. M. A. & Thornton, J. M. (2003). Structural characterisation and functional significance of transient protein-protein interactions. *Journal of Molecular Biology* **325**, 991-1018.
35. Robinson, C. V., Sali, A. & Baumeister, W. (2007). The molecular sociology of the cell. *Nature* **450**, 973-982.
36. Marsh, J. A., Rees, H. A., Ahnert, S. E. & Teichmann, S. A. (2015). Structural and evolutionary versatility in protein complexes with uneven stoichiometry. *Nature Communications* **6**, 1-10.
37. Zhou, Z. H. (2011). Atomic resolution cryo electron microscopy of macromolecular complexes. *Advances in Protein Chemistry and Structural Biology* **82**, 1-35.
38. Pieters, B. J. G. E., van Eldijk, M. B., Nolte, R. J. M. & Mecinovic, J. (2016). Natural supramolecular protein assemblies. *Chemical Society Reviews* **45**, 24-39.
39. Gelse, K., Pöschl, E. & Aigner, T. (2003). Collagens-structure, function, and biosynthesis. *Advanced Drug Delivery Reviews* **55**, 1531-1546.

40. Herrmann, H. & Aebi, U. (2004). Intermediate filaments: molecular structure, assembly mechanism, and integration into functionally distinct intracellular Scaffolds. *Annual Review in Biochemistry* **73**, 749-789.

41. Hingorani, M. M. & O'Donnell, M. (2000). A tale of toroids in DNA metabolism. *Nature Reviews Molecular Cell Biology* **1**, 22-30.

42. Chen, X., Antson, A. A., Yang, M., Li, P., Baumann, C., Dodson, E. J., Dodson, G. G. & Gollnick, P. (1999). Regulatory features of the trp operon and the crystal structure of the trp RNA-binding attenuation protein from Bacillus stearothermophilus. *Journal of Molecular Biology* **289**, 1003-1016.

43. Hall, A., Karplus, P. A. & Poole, L. B. (2009). Typical 2-cys peroxiredoxins: structures, mechanisms and functions. *The FEBS journal* **276**, 2469-2477.

44. Flohé, L. & Harris, J. R. (2007). Peroxiredoxin systems: structures and functions, *Springer science*.

45. Saibil, H. (2013). Chaperone machines for protein folding, unfolding and disaggregation. *Nature Reviews Molecular Cell Biology* **14**, 630-642.

46. Saccoccia, F., Di Micco, P., Boumis, G., Brunori, M., Koutris, I., Miele, A. E., Morea, V., Sriratana, P., Williams, D. L., Bellelli, A. & Angelucci, F. (2012). Moonlighting by different stressors: crystal structure of the chaperone species of a 2-cys peroxiredoxin. *Structure* **20**, 429-439.

47. Radjainia, M., Venugopal, H., Desfosses, A., Phillips, Amy J., Yewdall, N. A., Hampton, Mark B., Gerrard, Juliet A. & Mitra, Alok K. (2015). Cryo-electron microscopy structure of human peroxiredoxin-3 filament reveals the assembly of a putative chaperone. *Structure* **23**, 912-920.

48. Fromm, S. A., Bharat, T. A., Jakobi, A. J., Hagen, W. J. & Sachse, C. (2015). Seeing tobacco mosaic virus through direct electron detectors. *Journal of Structural Biology* **189**, 87-97.

49. Song, L., Hobaugh, M. R., Shustak, C., Cheley, S., Bayley, H. & Gouaux, J. E. (1996). Structure of staphylococcal α-hemolysin, a heptameric transmembrane pore. *Science* **274**, 1859-1866.

50. Virnau, P., Mallam, A. & Jackson, S. (2011). Structures and folding pathways of topologically knotted proteins. *Journal of Physics Condensed Matter* **23**, 33-101.

51. Tanaka, H., Kato, K., Yamashita, E., Sumizawa, T., Zhou, Y., Yao, M., Iwasaki, K., Yoshimura, M. & Tsukihara, T. (2009). The structure of rat liver vault at 3.5 angstrom resolution. *Science* **323**, 384-388.

52. McHugh, C. A., Fontana, J., Nemecek, D., Cheng, N., Aksyuk, A. A., Heymann, J. B., Winkler, D. C., Lam, A. S., Wall, J. S., Steven, A. C. & Hoiczyk, E. (2014). A virus capsid-like nanocompartment that stores iron and protects bacteria from oxidative stress. *EMBO Journal* **33**, 1896-1911.

53. Maji, S. K., Perrin, M. H., Sawaya, M. R., Jessberger, S., Vadodaria, K., Rissman, R. A., Singru, P. S., Nilsson, K. P., Simon, R., Schubert, D., Eisenberg, D., Rivier, J., Sawchenko, P., Vale, W. & Riek, R. (2009). Functional amyloids as natural storage of peptide hormones in pituitary secretory granules. *Science* **325**, 328-332.

54. Chiti, F. & Dobson, C. M. (2006). Protein misfolding, functional amyloid, and human disease. *Annual Review in Biochemistry* **75**, 333-366.

55. Hartl, F. U. & Hayer-Hartl, M. (2009). Converging concepts of protein folding *in vitro* and in vivo. *Nature Structural and Molecular Biology* **16**, 574-581.

56. Hamada, D. & Dobson, C. M. (2002). A kinetic study of β-lactoglobulin amyloid fibril formation promoted by urea. *Protein Science* **11**, 2417-2426.

57. Rambaran, R. N. & Serpell, L. C. (2008). Amyloid fibrils: abnormal protein assembly. *Prion* **2**, 112-117.

58. Harrison, R. S., Sharpe, P. C., Singh, Y. & Fairlie, D. P. (2007). Amyloid peptides and proteins in review. *Reviews of Physiology, Biochemistry and Pharmacology* **159**, 1-77.

59. Pedersen, J. S. & Otzen, D. E. (2008). Amyloid-a state in many guises: survival of the fittest fibril fold. *Protein Science* **17**, 2-10.

60. Fowler, D. M., Koulov, A. V., Balch, W. E. & Kelly, J. W. (2007). Functional amyloid-from bacteria to humans. *Trends in Biochemical Sciences* **32**, 217-224.

61. Klug, A. (1999). The tobacco mosaic virus particle: structure and assembly. *Philosophical Transactions of the Royal Society of Londond B: Biological Sciences* **354**, 531-535.

62. Andersson, E. K., Bengtsson, C., Evans, M. L., Chorell, E., Sellstedt, M., Lindgren, A. E. G., Hufnagel, D. A., Bhattacharya, M., Tessier, P. M., Wittung-Stafshede, P., Almqvist, F. & Chapman, M. R. (2013). Modulation of curli assembly and pellicle biofilm formation by chemical and protein chaperones. *Chemistry & Biology* **20**, 1245-1254.

63. Bokhove, M., Claessen, D., de Jong, W., Dijkhuizen, L., Boekema, E. J. & Oostergetel, G. T. (2013). Chaplins of streptomyces coelicolor self-assemble into two distinct functional amyloids. *Journal of Structural Biology* **184**, 301-309.

64. Kwan, A. H., Winefield, R. D., Sunde, M., Matthews, J. M., Haverkamp, R. G., Templeton, M. D. & Mackay, J. P. (2006). Structural basis for rodlet assembly in fungal hydrophobins. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 3621-3626.

65. Bayro, M. J., Daviso, E., Belenky, M., Griffin, R. G. & Herzfeld, J. (2012). An amyloid organelle, solid-state NMR evidence for cross-β assembly of gas vesicles. *Journal of Biological Chemistry* **287**, 3479-3484.

66. Iconomidou, V. A., Vriend, G. & Hamodrakas, S. J. (2000). Amyloids protect the silkmoth oocyte and embryo. *FEBS Letters* **479**, 141-145.

67. Keten, S. & Buehler, M. J. (2010). Nanostructure and molecular mechanics of spider dragline silk protein assemblies. *Journal of The Royal Society Interface* **7**, 1709-1721.

68. Fowler, D. M., Koulov, A. V., Alory-Jost, C., Marks, M. S., Balch, W. E. & Kelly, J. W. (2005). Functional amyloid formation within mammalian tissue. *PLoS Biology* **4**, 100-107.

69. Yoon, S. & Welsh, W. J. (2004). Detecting hidden sequence propensity for amyloid fibril formation. *Protein Science* **13**, 2149-2160.

70. Liang, Y., Lynn, D. G. & Berland, K. M. (2010). Direct observation of nucleation and growth in amyloid self-assembly. *Journal of the American Chemical Society* **132**, 6306-6308.

71. Naiki, H., Hashimoto, N., Suzuki, S., Kimura, H., Nakakuki, K. & Gejyo, F. (1997). Establishment of a kinetic model of dialysis-related amyloid fibril extension *in vitro*. *Amyloid* **4**, 223-232.

72. Serio, T. R., Cashikar, A. G., Kowal, A. S., Sawicki, G. J., Moslehi, J. J., Serpell, L., Arnsdorf, M. F. & Lindquist, S. L. (2000). Nucleated conformational conversion and the replication of conformational information by a prion determinant. *Science* **289**, 1317-21.

73. Lee, J., Culyba, E. K., Powers, E. T. & Kelly, J. W. (2011). Amyloid-β forms fibrils by nucleated conformational conversion of oligomers. *Nature Chemical Biology* **7**, 602-609.

74. Schmit, J. D., Ghosh, K. & Dill, K. (2011). What drives amyloid molecules to assemble into oligomers and fibrils? *Biophysical Journal* **100**, 450-458.

75. Cohen, A. S. & Calkins, E. (1959). Electron microscopic observations on a Fibrous Component in Amyloid of Diverse Origins. *Nature* **183**, 1202-1203.

76. Eanes, E. D. & Glenner, G. G. (1968). X-ray differaction styudies on amyloid filaments. *Journal of Histochemistry and Cytochemistry* **16**, 673-677.

77. Jahn, T. R., Makin, O. S., Morris, K. L., Marshall, K. E., Tian, P., Sikorski, P. & Serpell, L. C. (2010). The common architecture of Cross-β amyloid. *Journal of Molecular Biology* **395**, 717-727.

78. Sunde, M., Serpell, L. C., Bartlam, M., Fraser, P. E., Pepys, M. B. & Blake, C. C. (1997). Common core structure of amyloid fibrils by synchrotron X-ray diffraction. *Journal of Molecular Biology* **273**, 729-739.

79. Serpell, L. C. & Smith, J. M. (2000). Direct visualisation of the β-sheet structure of synthetic Alzheimer's amyloid. *Journal of Molecular Biology* **299**, 225-231.

80. Serpell, L. C., Sunde, M., Benson, M. D., Tennent, G. A., Pepys, M. B. & Fraser, P. E. (2000). The protofilament substructure of amyloid fibrils. *Journal of Molecular Biology* **300**, 1033-1039.

81. Fändrich, M., Meinhardt, J. & Grigorieff, N. (2009). Structural polymorphism of Alzheimer Aβ and other amyloid fibrils. *Prion* **3**, 89-93.

82. Schmidt, M., Sachse, C., Richter, W., Xu, C., Fandrich, M. & Grigorieff, N. (2009). Comparison of Alzheimer Aβ(1-40) and Aβ(1-42) amyloid fibrils reveals similar protofilament structures. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 19813-19818.

83. Tycko, R. (2011). Solid state NMR studies of amyloid fibril structure. *Annual Review of Physical Chemistry* **62**, 279-299.

84. Nelson, R., Sawaya, M. R., Balbirnie, M., Madsen, A. O., Riekel, C., Grothe, R. & Eisenberg, D. (2005). Structure of the cross-β spine of amyloid-like fibrils. *Nature* **435**, 773-778.

85. Sawaya, M. R., Sambashivan, S., Nelson, R., Ivanova, M. I., Sievers, S. A., Apostol, M. I., Thompson, M. J., Balbirnie, M., Wiltzius, J. J., McFarlane, H. T., Madsen, A. O., Riekel, C. & Eisenberg, D. (2007). Atomic structures of amyloid cross-β spines reveal varied steric zippers. *Nature* **447**, 453-457.

86. Bertini, I., Gonnelli, L., Luchinat, C., Mao, J. & Nesi, A. (2011). A new structural model of Aβ40 fibrils. *Journal of the American Chemical Society* **133**, 16013-16022.

87. Goldsbury, C. S., Cooper, G. J., Goldie, K. N., Muller, S. A., Saafi, E. L., Gruijters, W. T., Misur, M. P., Engel, A., Aebi, U. & Kistler, J. (1997). Polymorphic fibrillar assembly of human amylin. *Journal of Structural Biology* **119**, 17-27.

88. Diaz-Avalos, R., Long, C., Fontano, E., Balbirnie, M., Grothe, R., Eisenberg, D. & Caspar, D. L. (2003). Cross-β order and diversity in nanocrystals of an amyloid-forming peptide. *Journal of Molecular Biology* **330**, 1165-1175.

89. Petkova, A. T., Leapman, R. D., Guo, Z., Yau, W. M., Mattson, M. P. & Tycko, R. (2005). Self-propagating, molecular-level polymorphism in Alzheimer's β-amyloid fibrils. *Science* **307**, 262-265.

90. Fitzpatrick, A. W. P., Debelouchina, G. T., Bayro, M. J., Clare, D. K., Caporini, M. A., Bajaj, V. S., Jaroniec, C. P., Wang, L., Ladizhansky, V., Müller, S. A., MacPhee, C. E., Waudby, C. A., Mott, H. R., De Simone, A., Knowles, T. P. J., Saibil, H. R., Vendruscolo, M., Orlova, E. V., Griffin, R. G. & Dobson, C. M. (2013). Atomic structure and hierarchical assembly of a cross-β amyloid fibril. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 5468-5473.

91. Otzen, D. (2010). Functional amyloid: turning swords into plowshares. *Prion* **4**, 256-264.

92. Gazit, E. (2007). Plenty of Room for Biology at the Bottom: An introduction to bionanotechnology. *Imperial College Press.*

93. Sivakumar, P. M., Kodolov, V. I., Zaikov, G. E. & Haghi, A. K. (2013). Nanostructure, nanosystems, and nanostructured materials: theory, production and development. *Apple Academic Press*.

94. Gazit, E. & Nussinov, R. (2008). Nanostructure design: methods and protocols, *Humana Press*.

95. Mandal, D., Nasrolahi Shirazi, A. & Parang, K. (2014). Self-assembly of peptides to nanostructures. *Organic & Biomolecular Chemistry* **12**, 3544-3561.

96. Vendruscolo, M. & Dobson, C. M. (2013). Structural biology: protein self-assembly intermediates. *Nature Chemical Biology* **9**, 216-217.

97. Gerrard, J. A. (2013). Protein nanotechnology: protocols, instrumentation, and applications. *Humana Press*.

98. Daraio, C. & Jin, S. (2012). Synthesis and patterning methods for nanostructures Useful for biological applications. In nanotechnology for biology and medicine: at the building block level. *Springer science.*

99. Bhatia, D., Sharma, S. & Krishnan, Y. (2011). Synthetic, biofunctional nucleic acid-based molecular devices. *Current Opinion in Biotechnology* **22**, 475-84.

100. Krishnan, Y. & Simmel, F. C. (2011). Nucleic acid based molecular devices. *Angewandte Chemie International Edition in English* **50**, 3124-3156.

101. Namiki, Y., Fuchigami, T., Tada, N., Kawamura, R., Matsunuma, S., Kitamoto, Y. & Nakagawa, M. (2011). Nanomedicine for cancer: lipid-based nanostructures for drug delivery and monitoring. *Accounts of Chemical Research* **44**, 1080-1093.

102. Mahler, A., Reches, M., Rechter, M., Cohen, S. & Gazit, E. (2006). Rigid,self-assembled hydrogel composed of a modified aromatic dipeptide. *Advanced Materials* **18**, 1365-1370.

103. Domigan, L. J. (2013). Proteins and peptides as biological nanowires: towards biosensing devices. *Methods in Molecular Biology* **996**, 131-152.

104. Sasso, L., Suei, S., Domigan, L., Healy, J., Nock, V., Williams, M. A. & Gerrard, J. A. (2014). Versatile multi-functionalization of protein nanofibrils for biosensor applications. *Nanoscale* **6**, 1629-1634.

105. Lakshmanan, A., Zhang, S. & Hauser, C. A. (2012). Short self-assembling peptides as building blocks for modern nanodevices. *Trends in Biotechnology* **30**, 155-165.

106. Saven, J. G. (2010). Computational protein design: Advances in the design and redesign of biomolecular nanostructures. *Current Opinion in Colloid and Interface Science* **15**, 13-17.

107. King, N. P. & Lai, Y.-T. (2013). Practical approaches to designing novel protein assemblies. *Current Opinion in Structural Biology* **23**, 632-638.

108. King, N. P., Bale, J. B., Sheffler, W., McNamara, D. E., Gonen, S., Gonen, T., Yeates, T. O. & Baker, D. (2014). Accurate design of co-assembling multi-component protein nanomaterials. *Nature* **510**, 103-108.

109. Matthews, J. M. (2012). Protein dimerization and oligomerization in biology. *Springer Science*.

110. Heddle, J. G., Yokoyama, T., Yamashita, I., Park, S. Y. & Tame, J. R. (2006). Rounding up: engineering 12-membered rings from the cyclic 11-mer TRAP. *Structure* **14**, 925-933.

111. Miranda, F. F., Iwasaki, K., Akashi, S., Sumitomo, K., Kobayashi, M., Yamashita, I., Tame, J. R. & Heddle, J. G. (2009). A self-assembled protein nanotube with high aspect ratio. *Small* **5**, 2077-2084.

112. Ardini, M., Giansanti, F., Di Leandro, L., Pitari, G., Cimini, A., Ottaviano, L., Donarelli, M., Santucci, S., Angelucci, F. & Ippoliti, R. (2014). Metal-induced self-assembly of peroxiredoxin as a tool for sorting ultrasmall gold nanoparticles into one-dimensional clusters. *Nanoscale* **6**, 8052-8061.

113. Qiao, S. P., Lang, C., Wang, R. D., Li, X. M., Yan, T. F., Pan, T. Z., Zhao, L. L., Fan, X. T., Zhang, X., Hou, C. X., Luo, Q., Xu, J. Y. & Liu, J. Q. (2016). Metal induced self-assembly of designed V-shape protein into 2-D wavy supramolecular nanostructure. *Nanoscale* **8**, 333-341.

114. Yuan, J., Shen, J. & Kang, I.-K. (2008). Fabrication of protein-doped PLA composite nanofibrous scaffolds for tissue engineering. *Polymer International* **57**, 1188-1193.

115. Aluigi, A., Varesano, A., Montarsolo, A., Vineis, C., Ferrero, F., Mazzuchetti, G. & Tonin, C. (2007). Electrospinning of keratin/poly(ethylene oxide)blend nanofibers. *Journal of Applied Polymer Science* **104**, 863-870.

116. Claudio Tonin, A. A., Alessio Varesano and Claudia Vineis. (2010). Nanofibers: keratin-based nanofibres. *InTech*.

117. Barone, J. R., Schmidt, W. F. & Liebner, C. F. E. (2005). Thermally processed keratin films. *Journal of Applied Polymer Science* **97**, 1644-1651.

118. Baek, D. H., Ki, C. S., Um, I. C. & Park, Y. H. (2007). Metal ion adsorbability of electrospun wool keratose/silk fibroin blend nanofiber mats. *Fibers and Polymers* **8**, 271-277.

119. Thompson, Z. S., Rijal, N. P., Jarvis, D., Edwards, A. & Bhattarai, N. (2016). Synthesis of keratin-based nanofiber for biomedical engineering. *Journal of Visualized Experiments* **108**, 53381.

120. Valery, C., Artzner, F. & Paternostre, M. (2011). Peptide nanotubes: molecular organisations, self-assembly mechanisms and applications. *Soft Matter* **7**, 9583-9594.

121. Cinar, G., Ceylan, H., Urel, M., Erkal, T. S., Deniz Tekin, E., Tekinay, A. B., Dâna, A. & Guler, M. O. (2012). Amyloid Inspired self-sssembled peptide nanofibers. *Biomacromolecules* **13**, 3377-3387.

122. Robson Marsden, H. & Kros, A. (2009). Polymer-peptide block copolymers-an overview and assessment of synthesis methods. *Macromolecular Bioscience* **9**, 939-951.

123. Mason, J. M. & Arndt, K. M. (2004). Coiled coil domains: stability, specificity, and biological implications. *ChemBioChem* **5**, 170-176.

124. Castillo, J., Sasso, L. & Svendsen, W. E. (2012). Self-assembled peptide nanostructures: advances and applications in nanobiotechnology. *Pan Stanford publishing*.

125. Sawaya, M. R., Sambashivan, S., Nelson, R., Ivanova, M. I., Sievers, S. A., Apostol, M. I., Thompson, M. J., Balbirnie, M., Wiltzius, J. J. W., McFarlane, H. T., Madsen, A. O., Riekel, C. & Eisenberg, D. (2007). Atomic structures of amyloid cross-β spines reveal varied steric zippers. *Nature* **447**, 453-457.

126. Nilsson, M. R. (2004). Techniques to study amyloid fibril formation *in vitro*. *Methods* **34**, 151-60.

127. Gras, S. L., Waddington, L. J. & Goldie, K. N. (2011). Transmission electron microscopy of amyloid fibrils. *Methods in Molecular Biology* **752**, 197-214.

128. Langkilde, A. E., Morris, K. L., Serpell, L. C., Svergun, D. I. & Vestergaard, B. (2015). The architecture of amyloid-like peptide fibrils revealed by X-ray scattering, diffraction and electron microscopy. *Acta Crystallographica Section D: Biological Crystallography* **71**, 882-895.

129. Eisenberg, D. & Jucker, M. The amyloid state of Proteins in human diseases. *Cell* **148**, 1188-1203.

130. Aggeli, A., Bell, M., Boden, N., Keen, J. N., Knowles, P. F., McLeish, T. C., Pitkeathly, M. & Radford, S. E. (1997). Responsive gels formed by the spontaneous self-assembly of peptides into polymeric β-sheet tapes. *Nature* **386**, 259-262.

131. Aggeli, A., Nyrkova, I. A., Bell, M., Harding, R., Carrick, L., McLeish, T. C., Semenov, A. N. & Boden, N. (2001). Hierarchical self-assembly of chiral rod-like molecules as a model for peptide β-sheet tapes, ribbons, fibrils, and fibers. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 11857-11862.

132. Fenwick, R. B., Orellana, L., Esteban-Martín, S., Orozco, M. & Salvatella, X. (2014). Correlated motions are a fundamental property of β-sheets. *Nature Communications* **5**, 1-9.

133. Serpell, L. C. (2000). Alzheimer's amyloid fibrils: structure and assembly. *Biochimica et Biophysica Acta* **1502**, 16-30.

134. Valéry, C., Artzner, F., Robert, B., Gulick, T., Keller, G., Grabielle-Madelmont, C., Torres, M. L., Cherif-Cheikh, R. & Paternostre, M. (2004). Self-Association process of a peptide in solution: from β-Sheet filaments to large embedded Nanotubes. *Biophysical Journal* **86**, 2484-2501.

135. Wu, C., Scott, J. & Shea, J. E. (2012). Binding of Congo Red to amyloid protofibrils of the Alzhimer $A\beta_{9\text{-}40}$ peptide probed by molecular dynamics simulations. *Biophysical Journal* **103**, 550-557.

136. Krebs, M. R., Bromley, E. H. & Donald, A. M. (2005). The binding of thioflavin-T to amyloid fibrils: localisation and implications. *Journal of Structural Biology* **149**, 30-37.

137. Biancalana, M., Makabe, K., Koide, A. & Koide, S. (2009). Molecular mechanism of thioflavin-T binding to the surface of β-rich peptide self-assemblies. *Journal of Molecular Biology* **385**, 1052-1063.

138. Ivancic, V. A., Ekanayake, O. & Lazo, N. D. (2016). Binding modes of thioflavin-T on the surface of amyloid fibrils studied by NMR. *ChemPhysChem* **17**, 1-5.

139. Gorbenko, G., Trusova, V., Kirilova, E., Kirilov, G., Kalnina, I., Vasilev, A., Kaloyanova, S. & Deligeorgiev, T. (2010). New fluorescent probes for detection and characterization of amyloid fibrils. *Chemical Physics Letters* **495**, 275-279.

140. Huiyuan, L., Farid, R., Sharmistha, S., Panchanan, M., Gal, B., & Murakami, K. Amyloids and protein aggregation-analytical methods: Encyclopedia of Analytical Chemistry. *John Wiley & sons*.

141. Rice, S. B., Chan, C., Brown, S. C., Eschbach, P., Han, L., Ensor, D. S., Stefaniak, A. B., Bonevich, J., Vladár, A. E., Hight Walker, A. R., Zheng, J., Starnes, C., Stromberg, A., Ye, J. & Grulke, E. A. (2013). Particle size distributions by transmission electron microscopy: an interlaboratory comparison case study. *Metrologia* **50**, 663-678.

142. van Grondelle, W., Iglesias, C. L., Coll, E., Artzner, F., Paternostre, M., Lacombe, F., Cardus, M., Martinez, G., Montes, M., Cherif-Cheikh, R. & Valery, C. (2007). Spontaneous fibrillation of the native neuropeptide hormone Somatostatin-14. *Journal of Structural Biology* **160**, 211-223.

143. Valery, C., Paternostre, M., Robert, B., Gulik-Krzywicki, T., Narayanan, T., Dedieu, J. C., Keller, G., Torres, M. L., Cherif-Cheikh, R., Calvo, P. & Artzner, F. (2003). Biomimetic organization: octapeptide self-assembly into nanotubes of viral capsid-like dimension. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 10258-10262.

144. Caron, P. (2002). Somatuline autogel, a new formulation of lanreotide for the treatment of acromegalic patients. *Annales d'Endocrinologie* **63**, 2519-2524.

145. Tenidis, K., Waldner, M., Bernhagen, J., Fischle, W., Bergmann, M., Weber, M., Merkle, M. L., Voelter, W., Brunner, H. & Kapurniotu, A. (2000). Identification of a penta- and hexapeptide of islet amyloid polypeptide (IAPP) with amyloidogenic and cytotoxic properties. *Journal of Molecular Biology* **295**, 1055-1071.

146. Reches, M., Porat, Y. & Gazit, E. (2002). Amyloid fibril formation by pentapeptide and tetrapeptide fragments of human calcitonin. *Journal of Biological Chemistry* **277**, 35475-35480.

147. Reches, M. & Gazit, E. (2004). Amyloidogenic hexapeptide fragment of medin: homology to functional islet amyloid polypeptide fragments. *Amyloid* **11**, 81-89.

148. Gazit, E. (2002). A possible role for pi-stacking in the self-assembly of amyloid fibrils. *FASEB Journal* **16**, 77-83.

149. Lu, K., Jacob, J., Thiyagarajan, P., Conticello, V. P. & Lynn, D. G. (2003). Exploiting amyloid fibril lamination for nanotube self-assembly. *Journal of the American Chemical Society* **125**, 6391-6393.

150. Reches, M. & Gazit, E. (2003). Casting metal nanowires within discrete self-assembled peptide nanotubes. *Science* **300**, 625-627.

151. Krysmann, M. J., Castelletto, V., McKendrick, J. E., Clifton, L. A., I, W. H., Harris, P. J. & King, S. M. (2008). Self-assembly of peptide nanotubes in an organic solvent. *Langmuir* **24**, 8158-6812.

152. Sasso, L., Vedarethinam, I., Emneus, J., Svendsen, W. E. & Castillo-Leon, J. (2012). Self-assembled diphenylalanine nanowires for cellular studies and sensor applications. *Journal of Nanoscience and Nanotechnolgy* **12**, 3077-3083.

153. Ketchem, R., Hu, W. & Cross, T. (1993). High-resolution conformation of gramicidin A in a lipid bilayer by solid-state NMR. *Science* **261**, 1457-1460.

154. Bong, D. T., Clark, T. D., Granja, J. R. & Ghadiri, M. R. (2001). Self-assembling organic nanotubes. *Angewandte Chemie Intertional Edition in English* **40**, 988-1011.

155. Ghadiri, M. R., Granja, J. R. & Buehler, L. K. (1994). Artificial transmembrane ion channels from self-assembling peptide nanotubes. *Nature* **369**, 301-304.

156. Motiei, L., Rahimipour, S., Thayer, D. A., Wong, C.-H. & Ghadiri, M. R. (2009). Antibacterial cyclic D,L-α-glycopeptides. *Chemical Communications*, 3693-3695.

157. Santoso, S., Hwang, W., Hartman, H. & Zhang, S. (2002). Self-assembly of surfactant-like peptides with variable glycine tails to form nanotubes and nanovesicles. *Nano Letters* **2**, 687-691.

158. Yanlian, Y., Ulung, K., Xiumei, W., Horii, A., Yokoi, H. & Shuguang, Z. (2009). Designer self-assembling peptide nanomaterials. *Nano Today* **4**, 193-210.

159. Zhao, X., Pan, F., Xu, H., Yaseen, M., Shan, H., Hauser, C. A. E., Zhang, S. & Lu, J. R. (2010). Molecular self-assembly and applications of designer peptide amphiphiles. *Chemical Society Reviews* **39**, 3480-3498.

160. Matson, J. B., Zha, R. H. & Stupp, S. I. (2011). Peptide self-assembly for crafting functional biological materials. *Current Opinion in Solid State and Materials Science* **15**, 225-235.

161. Nieuwland, M., Ruizendaal, L., Brinkmann, A., Kroon-Batenburg, L., van Hest, J. C. M. & Lowik, D. W. P. M. (2013). A structural study of the self-assembly of a palmitoyl peptide amphiphile. *Faraday Discussions* **166**, 361-379.

162. Cheetham, A. G., Zhang, P., Lin, Y.-a., Lock, L. L. & Cui, H. (2013). Supramolecular nanostructures formed by anticancer drug assembly. *Journal of the American Chemical Society* **135**, 2907-2910.

163. Altman, M., Lee, P., Rich, A. & Zhang, S. (2000). Conformational behavior of ionic self-complementary peptides. *Protein Science* **9**, 1095-1105.

164. Zhang, S. (2002). Emerging biological materials through molecular self-assembly. *Biotechnology Advances* **20**, 321-339.

165. Aulisa, L., Dong, H. & Hartgerink, J. D. (2009). Self-assembly of multidomain Peptides: sequence aariation allows control over cross-linking and viscoelasticity. *Biomacromolecules* **10**, 2694-2698.

166. Chi, E. Y., Frey, S. L., Winans, A., Lam, K. L. H., Kjaer, K., Majewski, J. & Lee, K. Y. C. (2010). Amyloid-β fibrillogenesis seeded by interface-induced peptide misfolding and self-assembly. *Biophysical Journal* **98**, 2299-2308.

167. Makin, O. S., Atkins, E., Sikorski, P., Johansson, J. & Serpell, L. C. (2005). Molecular basis for amyloid fibril formation and stability. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 315-320.

168. Ghadiri, M. R., Granja, J. R., Milligan, R. A., McRee, D. E. & Khazanovich, N. (1993). Self-assembling organic nanotubes based on a cyclic peptide architecture. *Nature* **366**, 324-327.

169. Vauthey, S., Santoso, S., Gong, H., Watson, N. & Zhang, S. (2002). Molecular self-assembly of surfactant-like peptides to form nanotubes and nanovesicles. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 5355-5360.

170. Drin, G., Casella, J. F., Gautier, R., Boehmer, T., Schwartz, T. U. & Antonny, B. (2007). A general amphipathic α-helical motif for sensing membrane curvature. *Nature Structural and Molecular Biology* **14**, 138-146.

171. Kohn, W. D., Mant, C. T. & Hodges, R. S. (1997). α-helical protein assembly motifs. *Journal of Biological Chemistry* **272**, 2583-2586.

172. Crick, F. (1953). The packing of α-helices: simple coiled-coils. *Acta Crystallographica* **6**, 689-697.

173. Chou, K. C., Maggiora, G. M., Nemethy, G. & Scheraga, H. A. (1988). Energetics of the structure of the four α-helix bundle in proteins. *Proceedings of the National Academy of Sciences of the United States of America* **85**, 4295-4299.

174. Petka, W. A., Harden, J. L., McGrath, K. P., Wirtz, D. & Tirrell, D. A. (1998). Reversible hydrogels from self-assembling artificial proteins. *Science* **281**, 389-392.

175. Kojima, S., Kuriki, Y., Yoshida, T., Yazaki, K. & Miura, K.-i. (1997). Fibril formation by an amphipathic α-helix forming polypeptide produced by gene engineering. *Proceedings of the Japan Academy* **73**, 7-11.

176. Kajava, A. V., Potekhin, S. A., Corradin, G. & Leapman, R. D. (2004). Organization of designed nanofibrils assembled from α-helical peptides as determined by electron microscopy. *Journal of Peptide Science* **10**, 291-297.

177. Zimenkov, Y., Dublin, S. N., Ni, R., Tu, R. S., Breedveld, V., Apkarian, R. P. & Conticello, V. P. (2006). Rational design of a reversible pH-responsive switch for peptide self-assembly. *Journal of the American Chemical Society* **128**, 6770-6771.

178. Gribbon, C., Channon, K. J., Zhang, W., Banwell, E. F., Bromley, E. H., Chaudhuri, J. B., Oreffo, R. O. & Woolfson, D. N. (2008). Magicwand: a single, designed peptide that assembles to stable, ordered α-helical fibers. *Biochemistry* **47**, 10365-10371.

179. Wang, C., Stewart, R. J. & Kopecek, J. (1999). Hybrid hydrogels assembled from synthetic polymers and coiled-coil protein domains. *Nature* **397**, 417-20.

180. Deng, Y., Zheng, Q., Liu, J., Cheng, C.-S., Kallenbach, N. R. & Lu, M. (2007). Self-assembly of coiled-coil tetramers in the 1.40 Å structure of a leucine-zipper mutant. *Protein Science* **16**, 323-328.

181. Mittl, P. R., Deillon, C., Sargent, D., Liu, N., Klauser, S., Thomas, R. M., Gutte, B. & Grutter, M. G. (2000). The retro-GCN4 leucine zipper sequence forms a stable three-dimensional structure. *Proceedings of the National Academy of Sciences of the United States of America* **97**, 2562-2566.

182. O'Shea, E. K., Lumb, K. J. & Kim, P. S. (1993). Peptide velcro: design of a heterodimeric coiled coil. *Current Biology* **3**, 658-667.

183. Kotch, F. W. & Raines, R. T. (2006). Self-assembly of synthetic collagen triple helices. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 3028-3033.

184. Hulmes, D. J. (2002). Building collagen molecules, fibrils, and suprafibrillar structures. *Journal of Structural Biology* **137**, 2-10.

185. Nowak, A. P., Breedveld, V., Pakstis, L., Ozbas, B., Pine, D. J., Pochan, D. & Deming, T. J. (2002). Rapidly recovering hydrogel scaffolds from self-assembling diblock copolypeptide amphiphiles. *Nature* **417**, 424-8.

186. Woolfson, D. N., Bartlett, G. J., Bruning, M. & Thomson, A. R. (2012). New currency for old rope: from coiled coil assemblies to α-helical barrels. *Current Opinion in Structural Biology* **22**, 432-41.

187. Ali, M. H. & Imperiali, B. (2005). Protein oligomerization: how and why? *Bioorganic and Medicinal Chemistry* **13**, 5013-5020.

188. Stenkamp, R. E. (2001). Protein quaternary structure: symmetry patterns. *John Wiley & sons*.

189. Sheinerman, F. B., Norel, R. & Honig, B. (2000). Electrostatic aspects of protein-protein interactions. *Current Opinion in Structural Biology* **10**, 153-9.

190. Heifetz, A., Katchalski-Katzir, E. & Eisenstein, M. (2002). Electrostatics in protein-protein docking. *Protein Science* **11**, 571-87.

191. Vizcarra, C. L. & Mayo, S. L. (2005). Electrostatics in computational protein design. *Current Opinion in Chemical Biology* **9**, 622-6.

192. Zhang, S. (2003). Fabrication of novel biomaterials through molecular self-assembly. *Nature Biotechnology* **21**, 1171-1178.

2 Chapter two

Materials and methods

2.1 Material

2.1.1 Chemicals and reagents

Unless otherwise stated all chemicals, reagents, and solvents were obtained from Sigma-Aldrich, Invitrogen, BDH Laboratory supplies or Merck, and are of analytical grade. Precast, sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) gels and molecular weight protein standards were obtained from Invitrogen iLife technologies. Media for chromatography analysis was supplied as pre-packed columns from GE Life Sciences (Auckland, New Zealand). Milli-Q water supplied in-house by advance A10 water purification system (Merck, Millipore).

2.1.2 Peptides

Peptide sequences were custom manufactured and supplied by Mimotopes Pty Ltd (Sydney, Australia) as acetate salts with a purity of > 95% as tested by matrix-assisted laser desorption ionization (MALDI) mass spectrometry by the provider. The peptide acetate salts were obtained as lyophilized powders and stored at -20°C. Some peptide variants as stated in appendix II were supplied by Prof. Justin Hodgkiss (Victoria University), Wellington, New Zealand. Table 2.1 provides the list of peptide sequences studied in this thesis work.

Table 2.1: *Peptide sequences studied*

Peptide	Composition	Source protein
1	Ac-IKHLSVN-NH_2	*Bovine* peroxiredoxin 3 (*Bt* Prx3)
2	Ac-MHIRLSFN-NH_2	*Bovine* β-lactoglobulin (*Bt* β-LG)
3	Ac-HEFISTAH-NH_2	*Mycobacterium tuberculosis* diaminopimelate decarboxylase (DAPDC)
4	Ac-IHVVKAMR-NH_2	*E. coli* UmuD' protein
5	Ac-CYRGLLDSEDC-NH_2	*O. aeries* cuticular keratin type I
6	Ac-CYRRLLEGEEC-NH_2	*O. aeries* cuticular keratin type II
7	Ac-CEINTYRGLLDSEDC-NH_2	*O. aeries* cuticular keratin type I

2.1.3 Plasmids

The pET plasmid vector system was used for protein expression[1]. This vector has been reported to give high protein expression yields and to be sensitive to selective protease cleavage. The *pET-151/D-TOPO* plasmid vector and *pET22b* plasmid vector system[2,3] contained the *pBR322* origin of replication (ORI), the lac operator that allows induction of expression by addition of isopropyl β-D-1 thiogalactopyranoside (IPTG) (section 2.5.2) and the amp gene for resistance to ampicillin. The *pET151/D-TOPO* vector used for expression of human peroxiredoxin 3 (*h*Prx3) wild type and selective muteins, contained the N-terminal 6*Histidine tag, TEV cleavage site and the *bla* gene for ampicillin resistance and expression.

2.1.4 Bacterial strains

Different *E. coli* strains as detailed in table 2.2 were used during the research for plasmid propagation, cloning and protein expression[4].

Table 2.2 *Bacterial strains used for cloning and expression of protein*

Bacterial strain	**Description**
XL1-Blue (Stratagene)	Genotype: recA1, endA1, gyrA96, thi-1, hsdR17, supE44, relA1, lac [F′ proAB lacIq ZΔM15 Tn10 (Tetr)]
Rosetta "DE3" (Novogen)	Genotype – F- *ompT hsdS*B[$B^{-}mB^{-}$] *gal dcm* [DE3] pRARE [Cam^{R}])
BL21 "DE3" (Stratagene)	Genotype: F^{-} ompT, hsdS(rB^{-} mB^{-}), dcm^{+}, TetR, galλ(DE3), endA, Hte [argU proL Cam^{R}] [argU ileY leuW Strep/Specr])

Competent cells of each of the strains used were either purchased from Invitrogen Life Technologies (U.S.A) or made in-house (section 2.4.4.1). All the microbial work was carried out under aseptic conditions in a laminar flow hood or with an updraft generated by a flame.

XL1-Blue and *DH5α* cells were used for propagation of plasmid vectors. For expression of *h*Prx3 and mutant proteins, the *E. coli BL21 DE3 Rosetta* cells were used. The cells contain the gene for T7 RNA polymerase, required for expression from the *pET* plasmid vector, and are competent for high protein expression yield.

The cells contained the gene for resistance to chloramphenicol, to allow checking for non-specific bacterial growth in the media. The *BL21* (DE3) codon plus cells were used for expression of *rTEV* protein as they contained the *RILP* plasmid, allowing expression of rare codons, necessary for high-level expression of protein[1,5,6].

2.1.5 Sequences of protein investigated

The sequences of *h*Prx3 wild-type protein, an *h*Prx3 protein containing N-terminal 6*His tag and *h*Prx3 homodimer protein-β interface muteins are detailed in table 2.3 and 2.4.

Table 2.3: *Sequences of hPrx3 protein used in the research work*

Gene	Number of amino acids	Sequence
*h*Prx3	255	**MAAAVGRLLRASVARHVSAIPWGISATAALRPAACGRTSLTNL LCSGSSQAKLFSTSSSCH**APAVTQHAPYFKGTAVVNGEFKDLSLD DFKGKYLVLFFYPLDFTFVCPTEIVAFSDKANEFHDVNCEVVAVSV DSHFSHLAWINTPRKNGGLGHMNIALLSDLTKQISRDYGVLLEGSG LALRGLFIIDPNGVIKHLSVNDLPVGRSVEETLRLVKAFQYVETHGE VCPANWTPDSPTIKPSPAASKEYFQKVN
*h*Prx3 6*His^	227	**MHHHHHHGKPIPNPLLGLDSTENLYFQGIDPFT**APAVTQHAPYF KGTAVVNGEFKDLSLDDFKGKYLVLFFYPLDFTFVCPTEIVAFSDK ANEFHDVNCEVVAVSVDSHFSHLAWINTPRKNGGLGHMNIALLSD LTKQISRDYGVLLEGSGLALRGLFIIDPNGVIKHLSVNDLPVGRSVE ETLRLVKAFQYVETHGEVCPANWTPDSPTIKPSPAASKEYFQKVN
cleaved *h*Prx3	200	QGIDPFTAPAVTQHAPYFKGTAVVNGEFKDLSLDDFKGKYLVLFFY PLDFTFVCPTEIVAFSDKANEFHDVNCEVVAVSVDSHFSHLAWINT PRKNGGLGHMNIALLSDLTKQISRDYGVLLEGSGLALRGLFIIDPNG VIKHLSVNDLPVGRSVEETLRLVKAFQYVETHGEVCPANWTPDSPT IKPSPAASKEYFQKVN

Note: ***1)*** *The highlighted sequence in black corresponds to the mitochondrial localization sequence at the N-terminal.*

2) *^ is a truncated form of the protein with the mitochondrial localization sequence replaced with the peptide tag containing the six consecutive histidine residues (6*His).*

Table 2.4: *Sequences of hPrx3 muteins used in the research work*

Gene	Number of amino acids	Sequence
*h*Prx3 H136A	227	MHHHHHHGKPIPNPLLGLDSTENLYFQGIDPFTAPAVTQHAPYFKG TAVVNGEFKDLSLDDFKGKYLVLFFYPLDFTFVCPTEIVAFSDKAN EFHDVNCEVVAVSVDSHFSHLAWINTPRKNGGLGHMNIALLSDLT KQISRDYGVLLEGSGLALRGLFIIDPNGVIK**A**LSVNDLPVGRSVEET LRLVKAFQYVETHGEVCPANWTPDSPTIKPSPAASKEYFQKVN
*h*Prx3 H136Y ^	227	MHHHHHHGKPIPNPLLGLDSTENLYFQGIDPFTAPAVTQHAPYFKG TAVVNGEFKDLSLDDFKGKYLVLFFYPLDFTFVCPTEIVAFSDKAN EFHDVNCEVVAVSVDSHFSHLAWINTPRKNGGLGHMNIALLSDLT KQISRDYGVLLEGSGLALRGLFIIDPNGVIK**Y**LSVNDLPVGRSVEET LRLVKAFQYVETHGEVCPANWTPDSPTIKPSPAASKEYFQKVN
*h*Prx3 H136W	227	MHHHHHHGKPIPNPLLGLDSTENLYFQGIDPFTAPAVTQHAPYFKG TAVVNGEFKDLSLDDFKGKYLVLFFYPLDFTFVCPTEIVAFSDKAN EFHDVNCEVVAVSVDSHFSHLAWINTPRKNGGLGHMNIALLSDLT KQISRDYGVLLEGSGLALRGLFIIDPNGVIK**W**LSVNDLPVGRSVEET LRLVKAFQYVETHGEVCPANWTPDSPTIKPSPAASKEYFQKVN
*h*Prx3 H136F	227	MHHHHHHGKPIPNPLLGLDSTENLYFQGIDPFTAPAVTQHAPYFKG TAVVNGEFKDLSLDDFKGKYLVLFFYPLDFTFVCPTEIVAFSDKAN EFHDVNCEVVAVSVDSHFSHLAWINTPRKNGGLGHMNIALLSDLT KQISRDYGVLLEGSGLALRGLFIIDPNGVIK**F**LSVNDLPVGRSVEET LRLVKAFQYVETHGEVCPANWTPDSPTIKPSPAASKEYFQKVN

Note: *The highlighted residue in black corresponds to the mutated residue.*

2.1.6 Growth media

For propagation of plasmids and growth and protein expression from bacterial cells, the growth media detailed in table 2.5 were used. The media were sterilized by autoclaving and stored at room temperature before use[7].

Table 2.5: *Composition of media used for growth of bacterial cultures*

Medium	Composition
LB medium	Bacto-tryptone (10 g/L), yeast extract (5 g/L) and NaCl (10 g/L)
LB Agar medium	Bacto-tryptone (10 g/L), yeast extract (5 g/L), NaCl (10 g/L) and agar (13.7 g/L)
SOB medium	Bacto-tryptone (20 mg/mL), yeast extract (5 mg/mL), NaCl (0.58 g/L), KCl (0.95 g/L), $MgCl_2$ (1.20 g/L), $MgSO_4$ (1.20 g/L)
Autoinduction medium	Bacto-tryptone (10 g/L), yeast extract (5 g/L), $(NH_4)_2SO_4$ (3.3 g/L), KH_2PO_4 (7.1 g/L), glucose (0.5 g/L), α-lactose (2.0 g/L), $MgSO_4$ (0.15 g/L)

2.1.7 Antibiotics and additives

Antibiotics were used to select the transformed bacterial strains containing the recombinant plasmids with the desired gene of interest. The plasmid vectors used in the research contained the gene for ampicillin or kanamycin resistance. The *pLysS* plasmid in *E. coli BL21* bacterial expression strain contained a gene for chloramphenicol resistance[8]. The antibiotics were prepared as stock solutions: ampicillin and kanamycin in water and chloramphenicol in ethanol and stored in aliquots of 1 mL at -20°C until used (Table 2.6).

Table 2.6: *Details of antibiotics used for selective growth of transformed cells*

Antibiotic	Make	Stock concentration	Working concentration
Ampicillin (Amp)	Sigma-Aldrich	100 mg/mL	100 µg/mL
Chloramphenicol (Chl)	Merck-Millipore	30 mg/mL	30 µg/mL

Additive: Isopropyl β-D-1 thiogalactopyranoside (IPTG) was used to induce protein expression in bacterial expression systems from the vectors carrying the *lac* gene, which include the *pET* plasmid vector[9]. A stock solution of IPTG at a concentration of 400 mM in deionized water was added to the media at a 1:1000 dilution to give a working concentration of 0.4 mM.

2.2 Buffers and solutions

2.2.1 Solutions for CD spectroscopy and SAXS

Table 2.7: *Solvents used for C.D. spectroscopy and SAXS analysis*

Solvent	Make	Recipe
Methanol	Merck	Different percentages of solvent were prepared with water (v/v)
Ethanol	Merck	
Hexafluoro-2-propanol (HFIP)	Sigma-Aldrich	
Trifluoroacetic acid (TFA)		
Acetic acid		
100 mM Imidazole	BDH Analytical chemicals	0.68 mg imidazole dissolved in 50 mL Milli-Q water, pH adjusted to 8.5 and volume adjusted to 100 mL

Note: Milli-Q water used in the analysis was supplied from Milli-Q water purification system (Merck-Millipore).

2.2.2 Buffers and solutions for plasmid extraction

Table 2.8: *Buffers used for extraction of plasmid DNA from bacterial cells*

Process	Buffer	Composition
Extraction of plasmid DNA	Buffer S1	50 mM Tris-HCl, 10 mM EDTA, 100 µg/mL RNase A, pH 8.0
	Buffer S2	200 mM sodium hydroxide (NaOH), 1% sodium dodecyl sulfate (SDS)
	Buffer S3	4.2 M guanidinium hydrochloride (GuHCl), 900 mM potassium acetate, pH 4.80
	Wash buffer	10 mM Tris-HCl, 80% ethanol, pH 7.50
	Tris-EDTA (TE) buffer	10 mM Tris-HCl, 1 mM EDTA, pH 8.0

2.2.3 Buffers and solutions for protein purification

Extraction of protein from bacterial expression systems and further purification used specialized buffers that are known to fit for storage of biomolecules such as proteins. The buffer solutions and reagents used in the extraction and purification of *h*Prx3 protein were adapted from the previously established method by Phillips *et al.*[10]

Table 2.9: *Buffer solutions used for protein extraction and purification*

Process	Buffer	Composition
Immobilized metal affinity chromatography	Lysis buffer (B1)	20 mM Hepes, 150 mM NaCl, 30 mM imidazole pH 8.0
	Stripping buffer (B2)	20 mM sodium phosphate, 0.5 mM NaCl, 50 mM EDTA, pH 7.40
	Recharge buffer (B3)	0.1 M nickel sulfate
	Loading buffer (B4)	20 mM Hepes, 150 mM NaCl, 30 mM imidazole, pH 8.0
	Elution buffer (B5)	20 mM Hepes, 150 mM NaCl, 500 mM imidazole, pH 8.0
	Storage solution (B6)	20% ethanol
Gel filtration chromatography	Equilibration buffer (B7)	20 mM Hepes, 150 mM NaCl, pH 8.0
	Elution buffer (B8)	20 mM Hepes, 150 mM NaCl, pH 8.0
Purification of His-tag cleaved protein	Loading buffer (B9)	20 mM Hepes, 150 mM NaCl, 30 mM imidazole, pH 8.0
	Elution buffer (B10)	20 mM Hepes, 150 mM NaCl, pH 8.0

2.2.4 Buffers and solutions for SDS-PAGE

MES running buffer: A 20X stock solution was supplied by Invitrogen Life Technologies U.S.A. A 1:20 dilution was performed to make a 1X working concentration of the buffer that was used for gel electrophoresis.

Sample loading buffer: For1 mL of 2X sample loading buffer, 500 µL of sample loading buffer was diluted with 500 µL of Milli-Q water. The protein was diluted 1:1 with the sample loading dye sample before running on the gel.

2.2.5 Buffers and solutions for *rTEV* protease purification

The recombinant tobacco etch virus (*rTEV*) protease enzyme was used to cleave the 6*His-tag present at the N-terminus end of the recombinant protein. The buffers used for the extraction and purification of the *rTEV* protease are listed in the table below.

Table 2.10: *Buffer solutions used for rTEV extraction and purification*

Process	Buffer	Composition
Immobilized metal affinity chromatography	Lysis buffer	20 mM sodium phosphate, 500 mM NaCl, 30 mM Imidazole pH 8.0
	Elution buffer	20 mM sodium phosphate, 500 mM NaCl, 500 mM Imidazole, pH 8.0
	Storage solution	50 mM Tris-HCl, 0.5 mM EDTA, 1 mM DTT; pH 8.0

2.2.6 Buffers for peptide and protein mixing studies

Sodium citrate buffer: 20 mM sodium citrate, 150 mM NaCl, pH 4.0

MES buffer: 20 mM MES, 150 mM NaCl, pH 6.0

Sodium borate buffer: 20 mM sodium borate, 150 mM NaCl, pH 9.0

Sodium phosphate buffer: 20 mM sodium phosphate, 150 mM NaCl; pH 7.4

Tris-Cl buffer: 20 mM Tris-HCl, 150 mM sodium chloride; pH 7.4

Hepes buffer: 20 mM Hepes, 150 mM sodium chloride; pH 8.0

PBS buffer: 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, 2.7 mM KCl, 140mM NaCl pH 8.0

Note: For reducing conditions the buffers were prepared containing 3 mM TCEP.

2.2.7 Buffers and solutions for redox study of peptides

Sodium citrate buffer: 50 mM sodium citrate, 150 mM sodium chloride, pH 4.0

MES buffer: 50 mM MES, 150 mM sodium chloride, pH 5.0

Tris-Cl buffer: 20 mM Tris-HCl, 150 mM sodium chloride, pH 7.60

Sodium borate buffer: 50 mM sodium borate, 150 mM sodium chloride, pH 9.0

The above mentioned buffers were prepared with the addition of 100 mM hydrogen peroxide (H_2O_2) to study the self-assembly of the peptides under induced oxidizing conditions. Studies on the peptide sequences under reducing conditions were performed in the above-described buffers contained with 10 mM TCEP in the buffer.

2.2.8 Solution for ThT fluorescence assay

A stock solution of 500 μM Thioflavin-T (ThT) was used for ThT binding assays. To prepare the stock solution, 8.0 mg of ThT (abcam®, Cambridge (M.A.), U.S.A) was weighed and dissolved in 5 mL of solvent (Milli-Q water or 100 mM imidazole or 150 mM NaCl solution or Hepes pH 8.0 buffer). The stock solution was diluted to a working concentration of 50 μM. Unless otherwise stated, a 1:1 dilution of the peptide sample with 50 μM ThT was performed to obtain a desired concentration of the peptide and working ThT concentration of 25 μM.

2.2.9 Solution for TEM

Uranyl acetate (SPI-Chem™, U.S.A) was used as a negative staining dye for transmission electron microscopy (TEM) analysis. To prepare a 2% working solution of uranyl acetate: 20 mg of uranyl acetate was dissolved in 10 mL of Milli-Q water and filtered through a 0.22 µm filter. The uranyl acetate solution was stored at 4°C in a small glass bottle covered with aluminum foil. Uranyl acetate is a radioactive and photosensitive staining agent. Therefore appropriate care was taken while making, storing and using the sample[11].

2.2.10 Buffers and solutions for mass spectrometry

Analysis solution: 30 mg ferulic acid dissolved in 0.5 mL 2-propanol and 0.5 mL of 0.5% triflouroacetic acid (TFA).

2.3 Chromatography instrument, column, and media

Chromatographic procedures were performed on the AKTA Explorer system (GE Life Sciences). Hi-Trap chelating columns (1 mL, GE Lifesciences) and His-Trap columns (5 mL, GE Life Sciences) were used for purification of recombinant proteins by IMAC recharged with a metal ion (section 2.5.5.4). For size exclusion chromatography, Superdex HiLoad™ 16/600 200 prep grade (GE Lifesciences) column (120 mL) was used.

2.4 Molecular Biology

2.4.1 Commercial plasmids

Plasmids (section 2.1.3) containing customized gene sequences and the corresponding primers were designed using the site-directed mutagenesis tool available from GenScript Inc., New Jersey (U.S.A.). The designed plasmid expression vector with gene sequences shown in table 2.2 were then provided by Epoch Life Science Inc., Texas (U.S.A).

2.4.2 DNA sequencing

The accuracy of the mutagenesis was confirmed by the Sanger sequencing method[12]. The analysis was carried out at the School of Biological Sciences, University of Canterbury (N.Z.) using an Applied Biosystems 3130xl genetic analyzer.

2.4.3 DNA concentration

DNA yield was determined by measurement of U.V. absorbance at 260 nm wavelength[13]. The DNA concentration was determined using a Thermo Fischer Scientific Inc. Nanodrop spectrophotometer, with a built-in automated function, to calculate the concentration of the DNA present in the sample using the equation 2.1.

$$\text{Concentration} = (A * \varepsilon l/b) \qquad \text{equation 2.1}$$

$$\text{Concentration of DNA} = (A_{260}/\mathcal{E}_{260})*1 \qquad \text{equation 2.2}$$

The calculation of the DNA concentration was achieved by a modified Beer-Lambert equation, where c is the concentration in ng/μL, A is the absorbance in AU, ε is the wavelength-dependent extinction coefficient in ng-cm/μL, b is the path length in cm.

For calculation, A_{260} is the absorption at 260 nm in absorbance units (AU), ε_{260} is the wavelength-dependent extinction coefficient at 260 nm (~ 50 ng/cm/µL) 1 cm was the path length (0.1 cm in the NanoDrop 2000)[14]. As the molarity of polymers such as DNA does not depend on the length, mass per unit length was employed for the calculation.

2.4.4 Competent cells

2.4.4.1 Making chemically competent cells

A frozen glycerol stock of chemically competent cells was freshly plated onto an LB + agar plate using a flame sterilized nichrome wire. The plate was incubated overnight at 37°C. A single colony of bacterial cells was used to inoculate a 10 mL LB starter culture, incubated overnight at 37°C, shaking at 170 rpm. 2 mL of the overnight grown culture was used to inoculate further a 200 mL flask culture of LB, incubated at 37°C, shaking at 170 rpm. Cells were harvested at OD_{600} ~1 AU, by centrifuging at 6,500 rpm for 15 minutes (centrifuge maintained at 4°C) and suspended in 40 mL volume of 0.1 M calcium chloride ($CaCl_2$). The cells were divided into 100 µL aliquots, flash-frozen and stored at -80°C[15,16].

2.4.4.2 Transformation of chemically competent cells

A fresh aliquot of chemically competent cells was taken from -80°C storage and thawed gently on the ice. The plasmid DNA sample was also thawed gently, and 1 µL of the plasmid DNA was added to the competent cells, gently mixed using the pipette tip. The transformation mix was incubated on ice for 15 minutes and then subjected to heat shock by incubating at 42°C for 45 seconds. Following the incubation, the mix was immediately placed on ice for 2 minutes to cool down. 100 µL of pre-warmed (~ 37°C) super optimal broth (SOB) media was added to the mix and then incubated at 37°C for 60 minutes, shaking at 160 rpm to regenerate the cells.

The cells were then centrifuged at 6,500 rpm for 15 minutes and the supernatant removed. The cells were further suspended in 100 µL LB media and plated onto freshly prepared LB + agar plates containing selective antibiotics (section 2.1.7) and incubated at 37°C for 16 hours. Plates containing colonies of successfully transformed cells were stored at 4°C before use (expiry period is one week)[17].

2.4.5 Bacterial cell culture

The desired strain of bacterial cells was streaked onto a freshly prepared LB + agar plate containing the appropriate antibiotics (section 2.1.7) using a sterilized nichrome wire. The plated bacterial cells were incubated at 37°C for 16 hours. A single colony of the bacterial strain was used to inoculate a fresh 10 mL starter culture containing appropriate antibiotics, incubated overnight at 37°C shaking at 160 rpm[17]. The overnight grown starter culture was subsequently used to inoculate larger scale cultures (section 2.5.3).

2.4.6 Preparation of glycerol stocks

Glycerol stocks were used for long term storage of bacterial strains and cells transformed with recombinant plasmids containing the gene of interest for protein expression. 10 mL of starter culture was prepared as described in section 2.4.5 and 2 mL of this culture was used to inoculate a 200 mL flask culture of sterile LB. The culture was incubated at 37°C for 4 hours, to allow growth and taken off incubation once the OD_{600} reached 0.6 AU. The cultured cells were cooled down to room temperature, and 700 µL of the culture was added to a 2 mL cryovial containing 300 µL of 50% pre-autoclaved glycerol (final concentration of glycerol in the stock was 15%). Triplicate aliquots were vortexed and flash frozen by suspending the cryovial in liquid nitrogen and stored at -80°C[18].

2.5 Production of recombinant protein

2.5.1 Transformation and protein expression

The commercially available *BL21(DE3)* strain of *E. coli* cells was co-transformed with *pET 151-d-Topo* plasmid vector[9] (Epoch Lifesciences, U.S.A.) containing the gene encoding *h*Prx3 and with the *pRARE* plasmid[19]. Chemically competent *E. coli BL21 (DE3)* bacterial cells were thawed on ice and aliquots of 2 µL each of i) the *pET-151-d-Topo* plasmid containing the *h*Prx3 gene and ii) the *pRARE* plasmid were added. The transformation was performed as described in section 2.4.4.2. Glycerol stocks of the transformed cells were flash frozen using liquid nitrogen and stored at -80°C. The glycerol stock was used to inoculate expression cultures (section 2.5.2).

2.5.2 Small-scale expression of protein

To determine best conditions for expression of the recombinant protein: optimum incubation time, induction by isopropyl β-D-1-thiogalatopyranoside (IPTG) and temperature of growth, small-scale expression was trialed. Two different types of induction methods: IPTG mediated induction[20] and auto-induction[21], were used for the expression of recombinant proteins.

2.5.2.1 IPTG-induced expression of protein

A frozen glycerol stock of transformed cells (section 2.5.1) was freshly plated onto an LB + agar plate using a flame sterilized nichrome wire. The plate was incubated overnight at 37°C. A single colony of transformed cells grown on an LB + agar plate was used to inoculate 10 mL LB starter culture containing appropriate antibiotics. Suspension cultures were grown in 50 mL volume conical flasks to an OD_{600} between 0.4 to 0.6 AU. IPTG was added at 0.1 mM or 0.2 mM or 0.4 mM concentration[22].

To examine the effect of temperature on protein expression, the cultures were subsequently incubated at 26°C or 37°C, with shaking at 170 rpm. 500 µL of the culture were taken at different time points and centrifuged at 12,000 rpm for 5 minutes to obtain a cell pellet; the supernatant was discarded. The cells were suspended in 100 µL of SDS-PAGE buffer and sonicated to lyse the cells (70% amplitude, 3*5 minutes). The lysed cells were centrifuged again at 12,000 rpm for 5 minutes, and the supernatant separated. 10 µL of supernatant and 10 µL cell pellet suspended in SDS-PAGE buffer was analyzed by SDS-PAGE gel electrophoresis to assess protein expression (section 2.6.2).

2.5.2.2 Protein expression by autoinduction

A single colony of transformed cells was used to inoculate 10 mL of autoinduction media (section 2.1.6) and grown at 26°C and 37°C with shaking at 170 rpm[21]. Samples collected at different time points were analyzed as described in the above section.

2.5.3 Scale-up of protein expression

For expression of recombinant protein at a larger scale, LB + agar plates containing appropriate antibiotics were streaked with transformed cells and incubated overnight at 37°C. A single colony was picked to inoculate a 10 mL starter culture (with antibiotics) and incubated overnight with shaking at 170 rpm. The overnight grown culture was further used to inoculate flask cultures.

2.5.4 Protein expression by IPTG induction

The overnight grown starter culture was used to inoculate 200-500 mL of LB culture in a 2 L baffled flask containing an appropriate concentration of antibiotics. The flask culture was incubated at 37°C shaking at 170 rpm until the OD_{600} reached 0.4-0.6 AU.

To induce protein expression, the culture was brought to room temperature and induced with 0.4 mM IPTG. The induced culture was transferred to 26°C and grown overnight shaking at 170 rpm. Further cells were harvested by centrifugation at 6000 rpm for 15 minutes and re-suspended in 20 mL of lysis buffer. The cell suspension was directly used for extraction and purification of the protein or flash frozen using liquid nitrogen and stored at -80°C before use.

2.5.5 Purification of recombinant protein

2.5.5.1 Protein extraction by sonication

Lysis of cells was carried out using a sonicator 3000 (Misonix) equipped with a microprobe for low volume samples (1-5 mL) and a standard probe for larger volume (30 mL). A cell suspension was prepared by suspending the cell pellet in lysis buffer in a 50 mL plastic beaker. The beaker was placed in an ice slurry to prevent excessive heating of the sample. The cells were sonicated at 70% amplitude with a pulse rate of 2 seconds and 5 seconds off, six times for a total of 30 minutes.

2.5.5.2 Protein extraction by high-pressure homogenization

Cell lysis was achieved by passing the cell suspension through a valve and by applying a high pressure that causes disruption of cells by shear stress[23]. A 50 mL cell suspension was lysed by passing it through a Microfluidics M-110P cell disruptor at 17,000 psi maintained at 4°C. The cell suspension was passed through the homogenizer three times to achieve complete lysis. The lysed cells were centrifuged to separate the cell debris at 12,000 rpm for 30 minutes maintained at 4°C. The supernatant containing the soluble recombinant protein was subjected to SDS-PAGE gel electrophoresis to assess protein expression (section 2.6.2).

2.5.5.3 Concentration of protein

Protein samples were concentrated to the desired volume using centrifugal concentrators. Vivaspin® centrifugal concentrators (Sartorius Stedim Biotech) with a molecular weight cut-off (MWCO) of one-third of the molecular weight of the protein were selected. The Vivaspin® concentrator has two chambers: (i) sample chamber (top part) and (ii) flow through the collect chamber (bottom part). The protein sample was loaded into the top sample chamber and centrifuged at 4000 rpm and 4°C until the desired volume was obtained. 0.5 mL volume Vivaspin® centrifugal concentrators were used for small sample volumes.

2.5.5.4 Immobilized metal affinity chromatography (IMAC)

IMAC was used to purify the 6*His tagged recombinant protein using a pre-packed column (section 2.3). The clarified lysate was passed through a 0.45 µm filter and a 0.22 µm filter. The filtered lysate was applied to the 5 mL His-Trap column pre-equilibrated with 10 column volumes (CV) of lysis buffer (section 2.2.3). The column was washed with 5 column volumes of the lysis buffer to remove any unbound protein. The tagged protein was eluted from the column by applying a gradient of 100% IMAC elution buffer over 20 minutes. Both the steps of tagged recombinant protein binding and elution were carried out using the AKTA Explorer system (GE Lifesciences) maintained at 4°C and operated at 1 mL/min flow rate. The eluted protein was collected as 1 mL fractions and assessed by SDS-PAGE gel electrophoresis (section 2.6.2). Fractions that contained the tagged protein were pooled and stored at 4°C before use.

2.5.5.5 Size exclusion chromatography (SEC)

The pooled protein fractions were concentrated to ~ 2 mL (concentration of ~ 40 mg/mL) using 5 mL centricon tubes of 10 kDa molecular weight cut off (section 2.5.5.3).

The concentrated protein sample was filtered through a 0.22 µm syringe filter and loaded onto a Superdex 200 16/60 column, pre-equilibrated with 2 CVs of buffer B7 (section 2.2.3) attached to the AKTA FPLC Explorer system (GE Lifesciences). The chromatographic separation was performed at 4°C with the column operated at a flow rate of 0.5 mL/min using the elution buffer B8 (section 2.2.3) and size separated protein samples collected as 1 mL fractions on a 96 well sample plate using the auto-sampler attached to the AKTA system. The protein fractions were identified by UV absorbance at 280 nm and assessed by SDS-PAGE gel electrophoresis (section 2.6.2).

2.5.5.6 Storage of purified protein

Protein samples assessed by UV absorbance at 280 nm and SDS-PAGE gel electrophoresis were pooled and concentrated to 10 mg/mL. 200 µL aliquots in 2 mL cryovials with glycerol (15%) were flash frozen using liquid nitrogen and stored at -80 °C.

2.5.6 Removal of N-terminus 6*His-tag

2.5.6.1 Expression and purification of tobacco etch virus protease

rTEV protease was expressed in electrocompetent *E.coli BL21 (DE3)* and purified using a protocol based on the methods developed by Blommel & Fox[24].

2.5.6.2 *r*TEV assisted cleavage of His-tag protein

An aliquot of *rTEV* protease sample was taken from -80°C storage and thawed at room temperature. The recombinant *h*Prx3 protein expressed using the *pET151-d-Topo* vector contained a 26 residue linker peptide containing the "ENLYFQG" sequence (*rTEV* cleavage site) at the N-terminus end of the protein.

A ratio of 1 µg *r*TEV to 20 µg recombinant protein was used to cleave the 6*His-tag from the *h*Prx3 protein, incubated at 4°C for 12 hours. Samples were collected at T0, T6 and T12 hours and checked by SDS-PAGE gel electrophoresis for cleavage of the tag.

2.5.6.3 Purification of cleaved protein

The incubated sample mix (*rTEV* + protein) was loaded onto a gravity column containing a 2.5 mL bed volume of Cu^{2+} His-pure matrix (Sigma-Aldrich). The sample was mixed well with the column beads suspended in loading buffer B9 (section 2.2.3) using a pipette and allowed to sit for half an hour to allow the cleavage of 6*His-tag. The cleaved *h*Prx3 protein was purified using elution buffer B10 (section 2.2.3) by gravity chromatography. The elution step was performed a couple of times to recover most of the cleaved protein from the column matrix. The column matrix was further washed with high salt buffer B5 (section 2.2.3) to elute bound His-tag.

Samples of eluted protein, high salt buffer wash, and matrix beads were assessed by SDS-PAGE gel electrophoresis (section 2.6.2) for the efficacy of cleavage. The cleaved protein sample was concentrated to a volume of 2 mL and further passed through a 5 mL HisTrap™ column (GE Life Sciences) to remove any uncleaved tagged protein and cleaved tag. The HisTrap™ purified protein was collected and loaded onto a HiPrep™ 26/10 desalting column attached to the AKTA Explorer FPLC system and collected as 1 mL fractions. The fractions were assessed by SDS-PAGE gel electrophoresis and pooled together. The pooled protein sample was brought to a concentration of 1 mg/mL, and 200 µL aliquots of the protein were flash frozen in cryovials containing glycerol (15%)[25].

2.5.7 Buffer exchange

2.5.7.1 Buffer exchange by dialysis membrane

Protein samples were exchanged into desired solution conditions by transferring the protein sample into a dialysis tubing of 10,000 Da (MWCO). The tubing was placed in a beaker containing the required buffer solution with a typical sample volume to dialysis solution volume ratio of 1:2000. Dialysis was carried out by gently swirling the sample in the dialysis buffer to increase the rate of exchange and maintained at 4°C. The dialysis process was typically repeated to achieve a complete exchange.

2.5.7.2 Buffer exchange by dialysis device

To perform the buffer exchange of protein samples at lower volumes (< 1 mL) and minimize loss of sample, Slide-A-Lyzer™ MINI dialysis devices (Thermo Fisher Scientific) with an MWCO of 7000 kDa were used. A sample volume of 100-500 µL was pipetted into the dialysis cup and capped. The sample was submerged into the dialysis buffer at 4°C, regularly topped up and gently swirled to increase the dialysis exchange rate.

2.5.8 Preparation of peptide acetate solutions

Peptide acetate solutions were prepared by dissolving a precisely weighed amount of lyophilized peptide acetate powder in Milli-Q water/solvent/buffer solutions w/w (section 2.2.1, 2.2.6 and 2.2.7). The peptide acetate solutions were vortexed for 20 to 30 seconds to mix and equilibrated at 4°C or room temperature for 24 hours before analysis unless otherwise stated.

2.5.9 Mixing of *h*Prx3 protein with peptide acetates

The mixing of Prx3 peptide acetate with *h*Prx3 protein was carried out in Hepes, NaCl, TCEP, pH 8.0 buffer (section 2.2.6) unless otherwise stated. The *h*Prx3 protein stored at -80°C was thawed at room temperature and buffer exchanged by dialysis into the buffer overnight (section 2.5.7.2). The peptide acetates were prepared in the buffer as described in section 2.5.8 at different concentrations and equilibrated at 4°C for 24 hours before mixing unless otherwise stated. Calculated volume of peptide acetate solution was added to the *h*Prx3 protein sample such that the required final ratio of peptide acetate to protein is achieved (Table 2.11).

Table 2.11: *Ratio of peptide-protein mixes*

Ratio (peptide to *h*Prx3 protein toroid)	**Peptide concentration in mix**	**Molar concentration of peptide in mix**
15000:1	5.0 mg/mL	4.9 mM
7500:1	2.5 mg/mL	2.4 mM
3000:1	1.0 mg/mL	1.0 mM
1500:1	0.50 mg/mL	0.5 mM
500:1	0.17 mg/mL	0.17 mM
50:1	17 µg/mL	17 µM
1:1	0.34 µg/mL	0.33µM
1/50:1	6.8 ng/mL	6.6 nM

Note: The protein concentration was kept same at 0.1 mg/mL (0.33 µM) for all the conditions.

2.5.10 Oxidation and reduction of peptide acetates

The lyophilized peptide acetates were prepared in Milli-Q water or appropriate buffer solutions (section 2.2.7) and incubated for 24 hours or more at room temperature. To study the peptide sequences under oxidizing conditions: 10 mM peptide acetate in buffer was prepared, and 11 µL of 30% hydrogen peroxide (H_2O_2) was added to the peptide acetate solution (989 µL water/buffer + 11 µL of H_2O_2). To study the effect of reducing conditions, on peptide self-assembly, the oxidized peptide acetate solutions were dialyzed into a buffer containing 10 mM

TCEP solution overnight. Control samples were prepared to dissolve the peptide acetate in water/buffer without any reducing agent or oxidizing agent. Control samples for reducing conditions were prepared with peptide acetate in water/buffer containing 10 mM TCEP. All the samples were equilibrated at room temperature for over 24 hours before analysis unless otherwise stated.

2.6 Biophysical characterization

2.6.1 Determination of protein concentration

Purified protein samples were analyzed for UV absorbance at 280 nm using a Thermo Fischer Nanodrop 2000. For calculation of protein concentration, the Nanodrop 2000 was blanked with the *h*Prx3 storage buffer and 2 µL of the protein sample was placed on the sample tip and the absorbance at 280 nm of the sample recorded[26]. The theoretical extinction coefficient of the protein calculated from the amino acid sequence of the protein using the Expasy online bioinformatics server[27] was used for the calculation of the concentration using equation 2.3.

$$\text{Concentration of protein (mg/mL)} = (A_{280} / \varepsilon_{280)}) * 1 \qquad \textit{equation 2.3}$$

Where, A_{280} is the absorption of the protein sample at 280 nm in AU, ε_{280} is the extinction coefficient of the protein at 280 nm, and 1 is the path length (1 mm in the case of Nanodrop 2000).

2.6.2 SDS-PAGE analysis

Denaturation of the sample was carried out by the addition of 10 µL of loading dye containing SDS to 10 µL of the protein sample, and the mixture was heated at 95°C for 5 minutes.

For reducing SDS-PAGE gels, the sample loading dye was prepared with 5% β-mercaptoethanol or TCEP where indicated. A 10 or 15 well precast Bolt® 4-12% gel cassette ((Invitrogen Life Technologies) was assembled on the Bolt gel electrophoresis system (Invitrogen Life Technologies). 3 µL of Novex pre-stained molecular weight protein standard (Invitrogen Life Technologies) was loaded as the marker, and 5 µL of the denaturation mix was loaded into the sample wells.

The gel was run in 1 X MES buffer made from a premade 20 X buffer concentrate (Invitrogen Life Technologies) at a constant voltage of 165 V for 35 minutes. After completion of the electrophoresis run, the gel was removed from the cassette and washed with distilled water. The gel was stained with Coomassie simply blue stain (made in-house) using a three-step process: (i) the gel was placed in distilled water and microwave heated for 20 s to remove SDS present in the running buffer, (ii) stain was added to the gel and microwave heated for 20 s and then kept on a rocking table shaking at 50 rpm for 1 hour to allow optimum staining and (iii) gel was destained submerged in distilled water for 1 hour, kept on a rocking table. The gel was imaged on the BioRad Gel Doc XR$^+$ system using the Image Lab™ software.

2.6.3 SEC coupled with in-line static light scattering (SEC-SLS)

The protein sample purified through the analytical SEC column was passed through the in-line SLS system that detects the light scattered and measures the concentration of each fraction. To perform SEC-SLS analysis of protein samples, the Superdex 200 GL 10/300 column (GE Lifesciences) was connected to the Viscotek 302-040 Triple Detector DPC/SEC systems (ATA Scientific-Malvern Instruments Ltd.). Protein samples of known concentration were filtered through a 0.22 µm syringe filter with a 150 µL sample pipetted into an HPLC grade glass vial.

100 µL of the protein sample was injected into the analytical SEC column pre-equilibrated with 3 CVs of running buffer, operated at a flow rate of 0.5 mL/min and maintained at room temperature. Bovine serum albumin (BSA) was run along with the protein samples as a reference standard at the beginning and end of each sample sequence. The estimation of molecular weight distributions of the protein samples was done by calculation of the absolute molecular weight, radius of hydration and size distributions from the refractive index (RI), intrinsic viscosity and right angle light scattering (RALS) measurements against the reference standard BSA (Sigma) of molecular weight 66.4 kDa. The weight average molecular weight was calculated by the equation:

$$M = K'.\frac{(LS)}{(RI)} \qquad \textit{equation 2.4}$$

Where K' is a constant is calculated from the equation:

$$K' = \frac{K_{RI}}{K_{LS}\,(dn/dc)} \qquad \textit{equation 2.5}$$

K_{RI} and K_{LS} are instrumentation calibration constants, LS is the light scattering signal, RI is the refractive index signal and dn/dc is the refractive index increment (0.186 mL/g for globular proteins)[28].

2.6.4 Circular dichroism spectroscopy

Far-UV circular dichroism (CD) spectra were collected with a Jasco J-815 spectrometer using a 1.0 cm quartz cell at room temperature. Micromolar solutions for spectroscopic analysis were prepared by diluting the peptide stock solution (1 mg/mL) to a working concentration of 0.008 mg/mL. Stock solutions were prepared by dissolving 1 mg of the lyophilized peptide powder into 1.0 mL of Milli-Q of appropriate diluent (water or buffer or solvent) and equilibrated overnight before analysis. Spectroscopic blanks were Milli-Q water or buffer or solvent (section 2.2.1) used to prepare the peptide acetate solutions.

The samples were degassed to remove any air bubbles by placing them in a desiccator with vacuum applied for 5 minutes. All spectra were baseline corrected to remove effects attributed to scattering from aggregated particles. For protein analysis, stock solutions of protein samples were diluted to a working concentration of 0.1 mg/mL using the appropriate buffer before analysis.

2.6.5 Cross polarized light microscopy

Optimal microscopy experiments were carried out at room temperature (~ 20°C). Thin preparations of the peptide acetates in water prepared mounted on top of a glass slide and a coverslip placed to fix the peptide sample in place, carefully avoiding build-up of air bubbles. The sample slide was mounted on the Leica DM2700 P upright polarization microscope. The two polarizers were oriented at right angles to each other to achieve cross-polarization (± 45°). The preparations were imaged using the inline Leica EC3 high-speed digital color camera attached to the microscope and Leica laser microdissection (LMD) software[29].

2.6.6 ThT binding assay

ThT fluorescence assay was used to analyze the presence of extended β-sheet assemblies for the peptide dissolved in aqueous solutions[30]. Peptide samples were dissolved in an aqueous solution of ThT so that the final concentrations in each sample were 25 μM ThT and 0.25-30 mg/mL of peptide acetate. The peptide acetate solutions were prepared by dilution in water/buffer from a stock solution of 60-100 mg/mL and left to equilibrate at 4°C 24 hour unless otherwise stated. Fluorescence intensity was recorded for 100 μL aliquots placed into sample well plates using a Molecular Devices SpectraMax M5 fluorimeter equipped with a plate reader. Samples were excited at 440 nm, and fluorescence intensity was recorded at 482 nm, using a 25 μM ThT solution in water/buffer as a blank.

The fluorescence intensity values reported were average of three recordings obtained for triplicate sample preparations. The error bars displayed in the ThT binding plots were obtained from the standard deviation between three measurements of the sample.

2.6.7 Turbidity assay

Peptide samples prepared in solvents (section 2.2.1, 2.2.6 and 2.2.7) were analyzed at 500 nm post equilibration for 24 hours at 4°C. Absorbance measurements were performed at 500 nm for 100 µL aliquots placed into sample well plates using a Molecular Devices SpectraMax M5 fluorimeter equipped with a plate reader.

2.6.8 Transmission electron microscopy (TEM)

Aqueous peptide solutions (15 µL drops) were adsorbed on carbon-coated Formvar 200-mesh copper grids (ProSciTech, Australia), successively deposited on water (3 times) and 2% uranyl acetate solution for one minute each. Excess liquid was removed using filter paper, and the grid was left to air dry for a few hours before imaging. The grids were examined using an FEI Morgagni 268D transmission electron microscope operating at 80 kV, with magnifications up to 180k. Micrographs were captured with a SIS/Olympus megapixel III digital camera mounted on a phosphor screen. ImageJ version 1.3 was used to analyze the TEM micrographs and perform measurements on the nanostructures imaged[31]. To reduce high and low frequency noise components and enhance the difference between the background and sample, a 2 nm to 20 nm band-pass filter was applied to the raw micrographs. Measurement of the extended nanostructures was done using the analyze tool: the scale on the micrograph was set using the set scale tool to the nanometer. The measurement was performed using the measure tool by drawing the line across the length and width of the nanostructure[32].

2.6.9 Small angle X-ray scattering (SAXS)

Small angle X-ray scattering was carried out at the Australian Synchrotron Radiation Facility (AS, Clayton, Australia) on the SAXS/WAXS beamline operated at 12 KeV using sample detector distances of 0.5 to 1 m depending on the experiments. Samples of the peptide acetate were prepared in solvents (Table 1.1) at 5% and 10% concentration (w/w) at room temperature. The samples were then transferred into 1.2-1.3 mm glass capillaries (Mark-Röhrchen für röntgenographische Aufnahmen) supplied by Capillary Tube Supplied Ltd. Cornwall (U.K.).

Calculated amounts of the sample were pipetted at the top of the capillary and centrifuged at 4000 rpm for 2 minutes to get the sample to the bottom of the capillary tube. The top of the capillary was sealed by applying hot melted wax and subsequent air drying. The sample capillaries were introduced into a capillary holder accommodating 20 capillaries. The diffraction patterns were recorded for a reciprocal spacing q ($Å^{-1}$) from 0.02 to 2.1 $Å^{-1}$ using a Dectris-Pilatus detector (1 M, 170 mm x 170 mm, effective pixel size, 172 µm x 172 µm) and recorded via a CCD (chip charged-coupled device). The momentum transfer $q = 4\pi(\sin\theta/\lambda)$, where 2θ was the scattering angle and λ, the monochromatic X-ray beam wavelength of 0.15 nm was used[33].

The scattering profile of the solvent and buffer was recorded and used for background subtraction. Data reduction and subtraction were performed using the ScatterBrain software (Australian Synchrotron) software[34].

2.6.10 Fourier transform infrared spectroscopy (FTIR)

Attenuated total reflectance (ATR) Fourier transform infrared spectra were collected at the Australian Synchrotron infrared microspectroscopy beamline, Melbourne (AS, Clayton, Australia). The vibrations of the molecular bonds resulting from absorption of infrared radiation passed through the sample were recorded as a spectrum that represents absorption as a function of wavenumbers[35,36]. For analysis of peptide secondary structure conformation: the amide I vibrations[37,38] were analyzed for β-sheet content[39]. 10% (w/w) samples of the peptide acetates in water and a small pellet of KBr standard (used for background subtraction) were placed between two diamond windows of a micro compression cell (Thermo Fisher Scientific, Waltham, MA, USA). The samples were measured at 4 cm^{-1} resolution using a Bruker V80v FT-IR spectrometer combined with a Hyperion 2000 focal plane array IR microscope fitted with a liquid nitrogen cooled narrow band mercury cadmium telluride (MCT) detector. Data was collected in transmission mode with a microscope aperture defined measurement area of 5 µm x 5 µm on the sample in the mid-infrared spectral range between 4000-700 cm^{-1}.

The sample spectra obtained from an average of 64 scans and water signal was removed by subtraction of pure water spectrum. The Bruker OPUS 7.2 software was used for the spectral analysis and calculation of the second derivative using a 9 to 15 point Savitzky-Golay filter.

2.6.11 Dynamic light scattering (DLS)

DLS measurements were performed on a Malvern Zetasizer Nano Ne Laser operated at 173° backscatter. The Zetasizer software version 3.0 from Malvern was used to collect and analyze the data. Five hundred µL of sample were measured in a disposable single use polystyrene half-micro cuvettes (Thermo Fischer Scientific) with a path length of 10 mm, at a controlled temperature of 25°C.

For three replicates of the sample, 10 runs of 10 s each were performed with three repetitions. The intensity size distribution, Z-average diameter and polydispersity index (PDI) were obtained from the autocorrelation function using the "protein analysis mode" for both peptides and peptide-protein mix. The error bars displayed in the DLS graph were obtained from the standard deviation between three measurements of the sample and data presented as % intensity vs. size distribution histogram[40].

2.6.12 Analytical ultracentrifugation (AUC)

AUC analysis was performed using a Beckman Optima XL-A analytical ultracentrifuge. Sample volumes of 380 µL were loaded onto the analytical ultracentrifuge cells with a six-channel carbon-filled epon centerpieces, to give a column height of 1.83 mm. The samples were first centrifuged at 5000 rpm in a Beckman Optima XL-A analytical ultracentrifuge using a Beckman An60-Ti analytical rotor. Solute distributions were recorded using absorption optics at a wavelength of 280 nm. When the sedimentation equilibrium was reached, the rotor speed was increased to 30,000 rpm and rotor temperature maintained at 5°C. The buffer viscosity and frictional ratios were calculated using the Sednterp software[41].

Sample preparation: Samples of *h*Prx3 protein WT and muteins stored at -80°C were thawed at room temperature. The protein sample was dialyzed into Hepes, NaCl, pH 8.0 buffer with TCEP overnight at 4°C (section 2.5.7.2). The protein sample was centrifuged at 5000 rpm for 5 minutes and subsequently filtered through a 0.22 µ filter. The filtered samples were subjected to AUC described above. The protein-peptide mix prepared as described in section 2.5.9. The samples were centrifuged at 5000 rpm for 2-3 minutes before loading onto AUC cells.

2.7 References

1. Rosano, G. L. & Ceccarelli, E. A. (2014). Recombinant protein expression in Escherichia coli: advances and challenges. *Frontiers in Microbiology* **5**, 1-17.
2. Cabrita, L. D., Dai, W. & Bottomley, S. P. (2006). A family of *E. coli* expression vectors for laboratory scale and high throughput soluble protein production. *BMC Biotechnology* **6**, 1-8.
3. Sinah, N., Williams, C. A., Piper, R. C. & Shields, S. B. (2012). A set of dual promoter vectors for high throughput cloning, screening, and protein expression in eukaryotic and prokaryotic systems from a single plasmid. *BMC Biotechnology* **12**, 1-11.
4. Terpe, K. (2006). Overview of bacterial expression systems for heterologous protein production: from molecular and biochemical fundamentals to commercial systems. *Applied Microbiology and Biotechnology* **72**, 211-222.
5. Hakes, D. J. & Dixon, J. E. (1992). New vectors for high level expression of recombinant proteins in bacteria. *Analytical Biochemistry* **202**, 293-298.
6. Blommel, P. G., Becker, K. J., Duvnjak, P. & Fox, B. G. (2007). Enhanced bacterial protein expression during auto-induction obtained by alteration of lac repressor dosage and medium composition. *Biotechnology Progress* **23**, 585-598.
7. Atlas, R. M. (2010). Handbook of Microbiological Media. *CRC Press*.
8. Davanloo, P., Rosenberg, A. H., Dunn, J. J. & Studier, F. W. (1984). Cloning and expression of the gene for bacteriophage T7 RNA polymerase. *Proceedings of the National Academy of Sciences of the United States of America* **81**, 2035-2039.
9. Bolivar, F., Rodriguez, R. L., Greene, P. J., Betlach, M. C., Heyneker, H. L., Boyer, H. W., Crosa, J. H. & Falkow, S. (1977). Construction and characterization of new cloning vehicles. II. A multipurpose cloning system. *Gene* **2**, 95-113.
10. Phillips, A. J., Littlejohn, J., Yewdall, N. A., Zhu, T., Valery, C., Pearce, F. G., Mitra, A. K., Radjainia, M. & Gerrard, J. A. (2014). Peroxiredoxin is a versatile self-assembling tecton for protein nanotechnology. *Biomacromolecules* **15**, 1871-1881.
11. Booth, D. S., Avila-Sakar, A. & Cheng, Y. (2011). Visualizing proteins and macromolecular complexes by negative stain EM: from grid preparation to image acquisition. *Journal of Visualized Experiments* **58**, 1-8.
12. Sanger, F. & Coulson, A. R. (1975). A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *Journal of Molecular Biology* **94**, 441-448.

13. Sambrook, J. & Russell, D. W. (2001). Molecular cloning: A laboratory manual. *Cold Spring Harbor Laboratory Press*.
14. Barbas, C. F., Burton, D. R., Scott, J. K. & Silverman, G. J. (2007). Quantitation of DNA and RNA. *Cold Spring Harbor Protocols*.
15. Inoue, H., Nojima, H. & Okayama, H. (1990). High efficiency transformation of *E. coli* with plasmids. *Gene* **96**, 23-28.
16. Tu, Z., He, G., Li, K. X., Chen, M. J., Chang, J., Chen, L., Yao, Q., Liu, D. P., Ye, H., Shi, J. & Wu, X. (2005). An improved system for competent cell preparation and high efficiency plasmid transformation using different *E. coli* strains. *European Journal of Biotechnology* **8**, 114-120.
17. Wilson, K. & Walker, J. (2010). Principles and techniques of biochemistry and molecular biology. *Cambridge University Press*.
18. Hubálek, Z. (2003). Protectants used in the cryopreservation of microorganisms. *Cryobiology* **46**, 205-229.
19. Novy, R., Drott, D., Yaeger, K. & Mierendorf, R. (2001). Overcoming the codon bias of *E. coli* for enhanced protein expression. *inNovations* **12**, 1-3.
20. Graslund, S., Nordlund, P., Weigelt, J., Hallberg, B. M., Bray, J., Gileadi, O., Knapp, S., Oppermann, U., Arrowsmith, C., Hui, R., Ming, J., dhe-Paganon, S., Park, H. W., Savchenko, A., Yee, A., Edwards, A., Vincentelli, R., Cambillau, C., Kim, R., Kim, S. H., Rao, Z., Shi, Y., Terwilliger, T. C., Kim, C. Y., Hung, L. W., Waldo, G. S., Peleg, Y., Albeck, S., Unger, T., Dym, O., Prilusky, J., Sussman, J. L., Stevens, R. C., Lesley, S. A., Wilson, I. A., Joachimiak, A., Collart, F., Dementieva, I., Donnelly, M. I., Eschenfeldt, W. H., Kim, Y., Stols, L., Wu, R., Zhou, M., Burley, S. K., Emtage, J. S., Sauder, J. M., Thompson, D., Bain, K., Luz, J., Gheyi, T., Zhang, F., Atwell, S., Almo, S. C., Bonanno, J. B., Fiser, A., Swaminathan, S., Studier, F. W., Chance, M. R., Sali, A., Acton, T. B., Xiao, R., Zhao, L., Ma, L. C., Hunt, J. F., Tong, L., Cunningham, K., Inouye, M., Anderson, S., Janjua, H., Shastry, R., Ho, C. K., Wang, D., Wang, H., Jiang, M., Montelione, G. T., Stuart, D. I., Owens, R. J., Daenke, S., Schutz, A., Heinemann, U., Yokoyama, S., Bussow, K. & Gunsalus, K. C. (2008). Protein production and purification. *Nature Methods* **5**, 135-146.
21. Studier, F. W. (2005). Protein production by auto-induction in high density shaking cultures. *Protein Expression and Purification* **41**, 207-234.

22. Sivashanmugam, A., Murray, V., Cui, C., Zhang, Y., Wang, J. & Li, Q. (2009). Practical protocols for production of very high yields of recombinant proteins using *E. coli*. *Protein Science* **18**, 936-948.
23. Middelberg, A. J. (2000). Microbial cell disruption by high-pressure homogenization. In downstream processing of proteins. *Humana Press*.
24. Blommel, P. G. & Fox, B. G. (2007). A combined approach to improving large-scale production of tobacco etch virus protease. *Protein Expression and Purification* **55**, 53-68.
25. Koh, C. M. (2013). Storage of bacteria and yeast. *Methods in Enzymology* **533**, 15-21.
26. Grimsley, G. R. & Pace, C. N. (2004). Spectrophotometric determination of protein concentration. *Current Protocols in Protein Science*.
27. Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M., Appel, R. & Bairoch, A. (2005). Protein identification and analysis tools on the ExPASy server. In the proteomics protocols handbook. *Humana Press*.
28. Wen, J., Arakawa, T. & Philo, J. S. (1996). Size-exclusion chromatography with on-line light-scattering, absorbance, and refractive index detectors for studying proteins and their interactions. *Analytical Biochemistry* **240**, 155-166.
29. Clare, B. H. & Abbott, N. L. (2005). Orientations of nematic liquid crystals on surfaces presenting controlled densities of peptides: amplification of protein-peptide binding events. *Langmuir* **21**, 6451-6461.
30. Munishkina, L. A. & Fink, A. L. (2007). Fluorescence as a method to reveal structures and membrane-interactions of amyloidogenic proteins. *Biochimica et Biophysica Acta: Biomembranes* **1768**, 1862-1885.
31. Ludtke, S. J., Baldwin, P. R. & Chiu, W. (1999). EMAN: semiautomated software for high-resolution single-particle reconstructions. *Journal of Structural Biology* **128**, 82-97.
32. Rice, S. B., Chan, C., Brown, S. C., Eschbach, P., Han, L., Ensor, D. S., Stefaniak, A. B., Bonevich, J., Vladár, A. E., Hight Walker, A. R., Zheng, J., Starnes, C., Stromberg, A., Ye, J. & Grulke, E. A. (2013). Particle size distributions by transmission electron microscopy: an interlaboratory comparison case study. *Metrologia* **50**, 663-678.
33. Valery, C., Artzner, F., Robert, B., Gulick, T., Keller, G., Grabielle-Madelmont, C., Torres, M. L., Cherif-Cheikh, R. & Paternostre, M. (2004). Self-association process of a peptide in solution: from β-Sheet filaments to large embedded nanotubes. *Biophysical Journal* **86**, 2484-2501.

34. Kirby, N. M., Mudie, S. T., Hawley, A. M., Cookson, D. J., Mertens, H. D. T., Cowieson, N. & Samardzic-Boban, V. (2013). A low-background-intensity focusing small-angle X-ray scattering undulator beamline. *Journal of Applied Crystallography* **46**, 1670-1680.

35. Barth, A. (2007). Infrared spectroscopy of proteins. *Biochimica et Biophysica Acta: Bioenergetics* **1767**, 1073-1101.

36. Bakshi, K., Liyanage, M. R., Volkin, D. B. & Middaugh, C. R. (2014). Fourier transform infrared spectroscopy of peptides. *Methods in Mololecular Biology* **1088**, 255-269.

37. Bandekar, J. (1992). Amide modes and protein conformation. *Biochimica et Biophysica Acta* **1120**, 123-43.

38. Haris, P. I. & Chapman, D. (1995). The conformational analysis of peptides using fourier transform infrared spectroscopy. *Biopolymers* **37**, 251-263.

39. Kawasaki, T., Fujioka, J., Imai, T., Torigoe, K. & Tsukiyama, K. (2014). Mid-infrared free-electron laser tuned to the amide I band for converting insoluble amyloid-like protein fibrils into the soluble monomeric form. *Lasers in Medical Science* **29**, 1701-1707.

40. Lorber, B., Fischer, F., Bailly, M., Roy, H. & Kern, D. (2012). Protein analysis by dynamic light scattering: methods and techniques for students. *Biochemistry and Molecular Biology Education* **40**, 372-382.

41. Schuck, P., Perugini, M. A., Gonzales, N. R., Howlett, G. J. & Schubert, D. (2002). Size-distribution analysis of proteins by analytical ultracentrifugation: strategies and application to model systems. *Biophysical Journal* **82**, 1096-1111.

3 Chapter three

Self-assembly of peptides derived from protein-β interfaces

3.1 Introduction

Biomolecular building blocks such as peptides are promising as molecular tectons assembling into a range of different nanostructures in solution. Their small size and chemical diversity make them highly desirable for designing nanopatterns for different applications such as bioelectronics, biosensors, and medical technology[1]. The design of stable peptide tectons is challenging. Choosing a natural source for designing self-assembling sequences gives the advantage that some properties of the parent molecule can be retained.

Peptides were designed based on protein interfaces in an attempt to create peptide tectons that retained some of the useful self-assembling properties of the natural interface. Proteins have a natural ability to self-assemble into hierarchical complexes as part of their function[2]. Many such complexes can associate and dissociate in a reversible manner in response to a biological trigger[3]. Such assembly of proteins is driven through protein-protein interfaces that can interact with each other in a reversible manner[4]. High-resolution crystal structures of such protein interfaces show interface sequences in noncovalent close contacts. It was thus hypothesized that the corresponding isolated protein interface sequences could serve as a source of peptide sequences that can self-assemble independently while retaining some properties of the parent interface.

3.2 Design models for peptide tectons

Peptide tectons are simple, short sequences that can spontaneously self-assemble into well-ordered nanoscale structures[5]. The different peptide sequences known to self-assemble to date belong to a limited group of biologically inspired sequences, as described in chapter one. New sources of design will thus widen the library of self-assembling sequences and provide for specific downstream applications in bionanotechnology. To explore the protein-protein interface sequences as a design source for novel self-assembling tectons, simple protein homo-interfaces were chosen. The simple β-continuous protein interface is reported to be present in more than 10% of solved protein homo-oligomers[4,6]. Such interfaces exhibit two identical β-strands from the two interacting protein units arranged in an anti-parallel manner in noncovalent close contact.

A survey of more than eight hundred entries in the protein data bank (PDB) for protein homomers performed by Dr. Céline Valéry (R.M.I.T University; Victoria, Australia) was downsized to 20 structures based on the following unbiased criteria: (i) selection of native homo-oligomers with β-continuous interfaces, excluding structures with ligands and co-factors, to limit complexity of the set of interactions involved; (ii) excluding fragmented and complex interfaces, that were comprised of multiple interacting motifs; (iii) selection of only one representative interface from the families of proteins with high sequence similarity, to ensure sequence variety and a valid test of the hypothesis.

Statistical analysis performed by Dr. Valéry on the set of structures using the PDBe PISA online tool (EMBL) for selection of final sequences generated details of protein-protein interfaces (Figure 3.1).

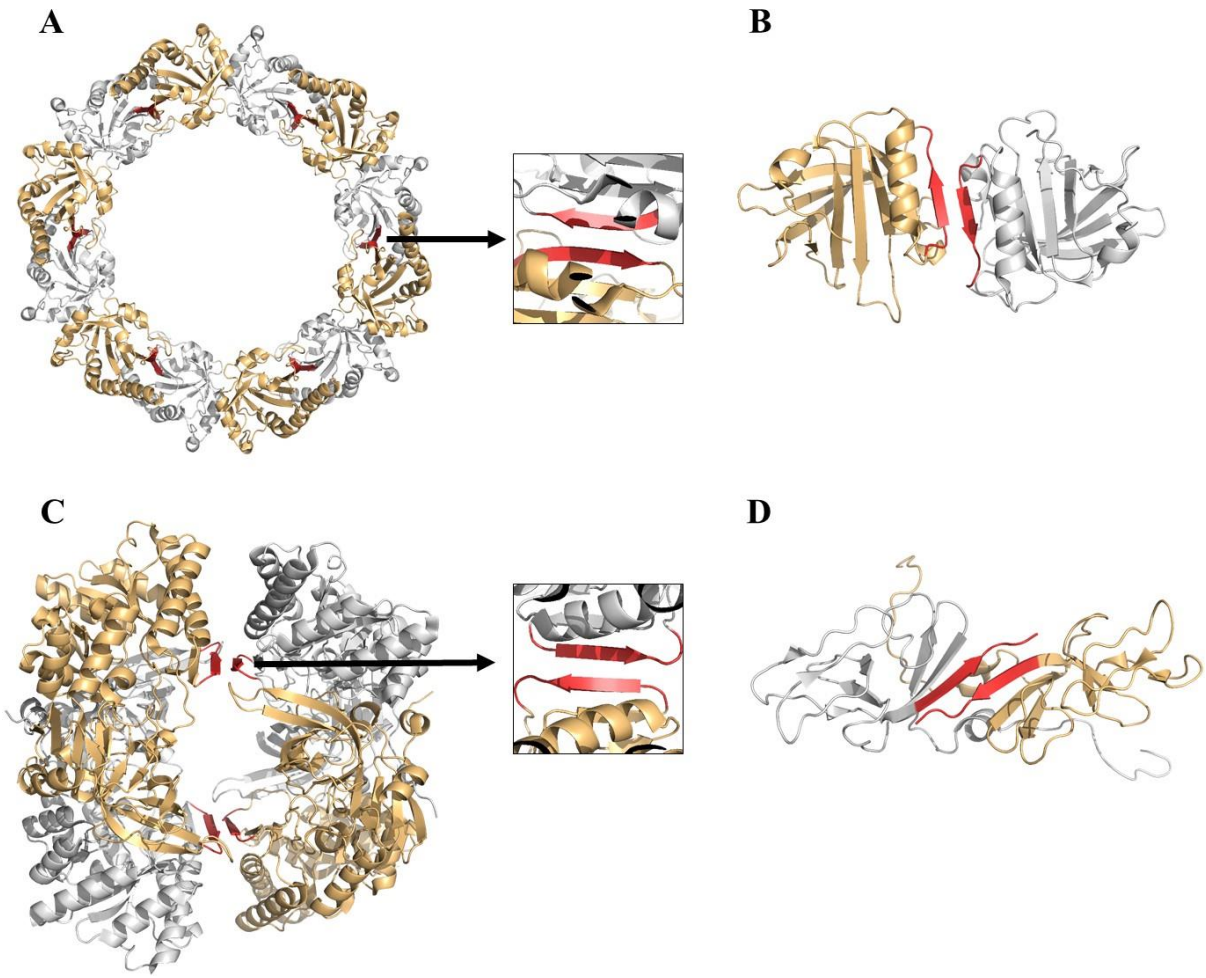


Figure 3.1: *Models of protein homo-oligomers: Crystal structures (X-ray crystallography) for A, B, and C and solution structure (NMR) for D. Red region highlights the interface contacts identified by the PISA software. (A) Bovine peroxiredoxin 3 (Prx3) homo-dodecamer (PDB: 1ZYE) (zoomed into the homodimer interface); (B) Bovine β-lactoglobulin (β-Lac) homodimer (PDB: 2Q39); (C) M. tuberculosis diaminopimelate decarboxylase (DAPDC) homo-tetramer (PDB: 2O0T) (zoomed into the tetramer interface); (D) E. coli UmuD' homodimer (PDB: 1I4V).*

Analysis of protein interface sequences showed that 90% of the sequences contained eleven or fewer residues, with 80% of them having at least one aromatic amino acid residue, almost all having 30% hydrophobic residues and 15% charged residues.

3.3 Peptide sequences investigated for self-assembly

The final set of selected peptide sequences were 7-8 residues long and corresponded to the β-continuous interface of the following protein homo-oligomers: (i) the bovine Prx3 homodimer within the dodecamer toroid[7]; (ii) the bovine β-Lac homodimer[8,9]; (iii) *M. tuberculosis* DAPDC homotetramer[10]; (iv) the *E.coli* UmuD' protein homodimer[11]. All the corresponding protein oligomers have been identified in solution, except for the *M. tuberculosis* DAPDC tetramer that is reported only in the crystal structure[12].The peptide sequences were capped at both ends: the N-terminus acetylated and the C-terminus amidated, to conserve the overall charge of the interface sequence. These peptides or similar sequences have not been earlier studied for self-assembly under any conditions, except for a β-Lac motif similar to the β-Lac peptide that was reported to self-assemble under reducing conditions in the presence of urea[13]. Table 3.1 presents the different design sources, corresponding β-interface sequences and their properties as calculated using the peptide property calculator {http://pepcalc.com/ (Innovogen, Sweden)}.

Table 3.1: *Protein β-interface sequences chosen as models*

Source protein and species	**PDB oligomer and code**	**Interface sequence (PISA)**	**Peptides**	**Theoretical pI**	**Net charge at pH 7**
Peroxiredoxin 3 *(bovine)*	1ZYE	IKHLSVN	Prx3 peptide	9.13	+ 1.0
β-Lactoglobulin *(bovine)*	2Q39, 2AKQ	MHIRLSFN	β-Lac peptide	9.52	+ 1.0
Diaminopimelate decarboxylase *(M. tuberculosis)*	2O0T	HEFISTAH	DAPDC peptide	5.98	-1.0
UmuD' *(E. coli)*	1I4V	IHVVKAMR	UmuD' peptide	11.0	+2.0

3.4 Characterization of peptide sequences in solution

Inspection of the peptide sequences using the peptide property calculator {http://pepcalc.com/ (Innovogen, Sweden)} showed that all of the four sequences have approximately 40% of residues that are hydrophobic. The distribution of hydrophobic residues and polar or charged residues within the peptide sequence highlighted their amphipathic nature that is consistent in all four peptide sequences. Amphipathic peptide sequences composed of alternating polar and non-polar residues show a strong tendency to self-assemble into extended amyloid-like fibrils[14,15].

The peptides were synthesized as a freeze-dried powder (acetate salts) to > 95% purity as judged by MALDI-TOF (Mimotopes Ltd., Sydney, Australia). Peptide acetate solutions were prepared by dissolving the lyophilized powder in Milli Q water (w/w) and incubated at 4°C for 24 hours (Section 2.5.8). The peptide acetates were soluble in water at dilute concentrations in the 0.25-2.5 mg/mL range for peptide 1 and peptide 2, ≥ 10 mg/mL for peptide 3 and ≥ 20 mg/mL for peptide 4. The pH of the peptide acetate samples recorded was ~ 5.0. Peptide acetates prepared in non-polar solvents: HFIP, TFA, and methanol at room temperature, however, were clear solutions with no residual particulate matter. This suggested that upon exposure to hydrophilic environments, the hydrophobic effects favor the burying of peptide surface to minimize solvent exposed surface area and drive peptide aggregation. This mechanism of self-assembly has been previously described for amphipathic β-sheet forming peptides and short peptides with a substantial number of hydrophobic residues in the sequence[14,16].

Turbidity assays were performed to assess the extent of peptide aggregation in solution. Peptide acetates incubated in water and polar solvents in the 10-40 mg/mL range were examined at 500

nm (section 2.6.7). Peptide acetates in water exhibited an increase in turbidity of the peptide solution, which can be attributed to the peptides self-assembled in solution[17,18]. No such change was, however, observed for the peptide acetates dissolved in HFIP or TFA or methanol.

Specific dyes such as ThT that show enhanced fluorescence on binding to amyloid-like assemblies provide a reliable way to test for confirmation and estimation of the specific assembly of peptides into extended β-sheet based structures[19]. Furthermore, ThT is soluble in water and many physiological buffers and does not interact with molten globule, unfolded states and amorphous aggregates of protein[20,21]. Thus, ThT binding assays provided a reliable way for detection and estimation of self-assembly of the peptides in solution. Samples of peptide acetates in water prepared at concentrations of 10-50 mg/mL mixed with ThT dye exhibited strong ThT binding (section 2.6.6), which suggested the presence of extended β-sheet based structures consistent in all the four cases (Figure 3.2).

For peptide acetate samples prepared by direct mixing of lyophilized peptide acetate powder with undiluted HFIP or TFA or methanol, no substantial ThT binding was observed. Conversely, peptide acetates prepared in dilute methanol solution (5%, 10% and 15% methanol in water) exhibited ThT binding, though to a lower extent than that observed for peptide acetates in water.

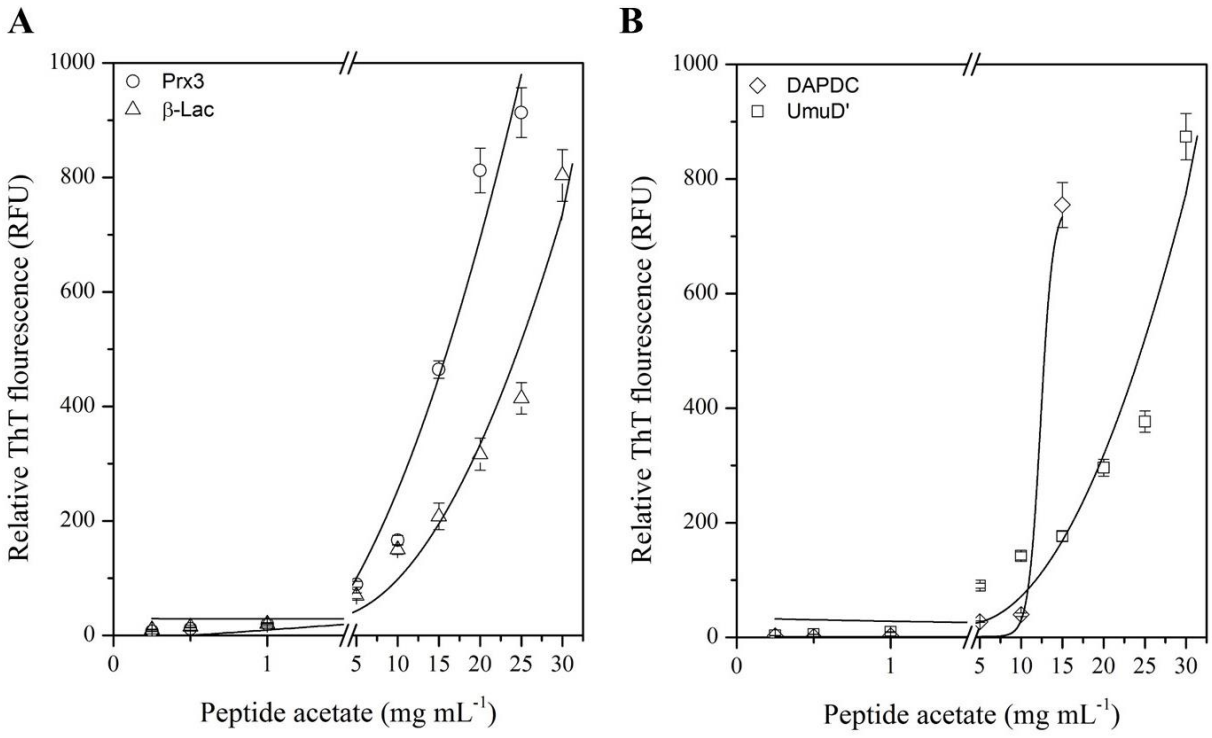


Figure 3.2: *ThT binding assays by dilution for peptide acetates: Fluorescence intensity values recorded immediately after dilution for Prx3 peptide, β-Lac and UmuD' peptides is fitted to the least squared linear function* $y = B + Ax + Ax^2$ *and the ThT data for DAPDC peptide is fitted to the sigmoid growth function* $y=B + (A + B)/(1 + e^{(x-x0)/dx})$*.The error bars correspond to the standard deviation between triplicate sample preparations.*

Samples of the peptide acetates were examined by transmission electron microscopy (TEM) to confirm fibril formation (section 2.6.8). TEM of negatively stained specimens allows visualization of extended assemblies such as amyloid-like fibrils at a high resolution in the dried state[22]. To enhance the resolution and contrast of the image for better quality, samples of the peptide acetates were centrifuged to pellet the peptide aggregates and subsequently treated with 3% glutaraldehyde[23,24]. TEM of peptide acetates in water adsorbed onto EM grids immediately after dilution revealed ordered nanostructures consistent for all four peptides (Figure 3.3).

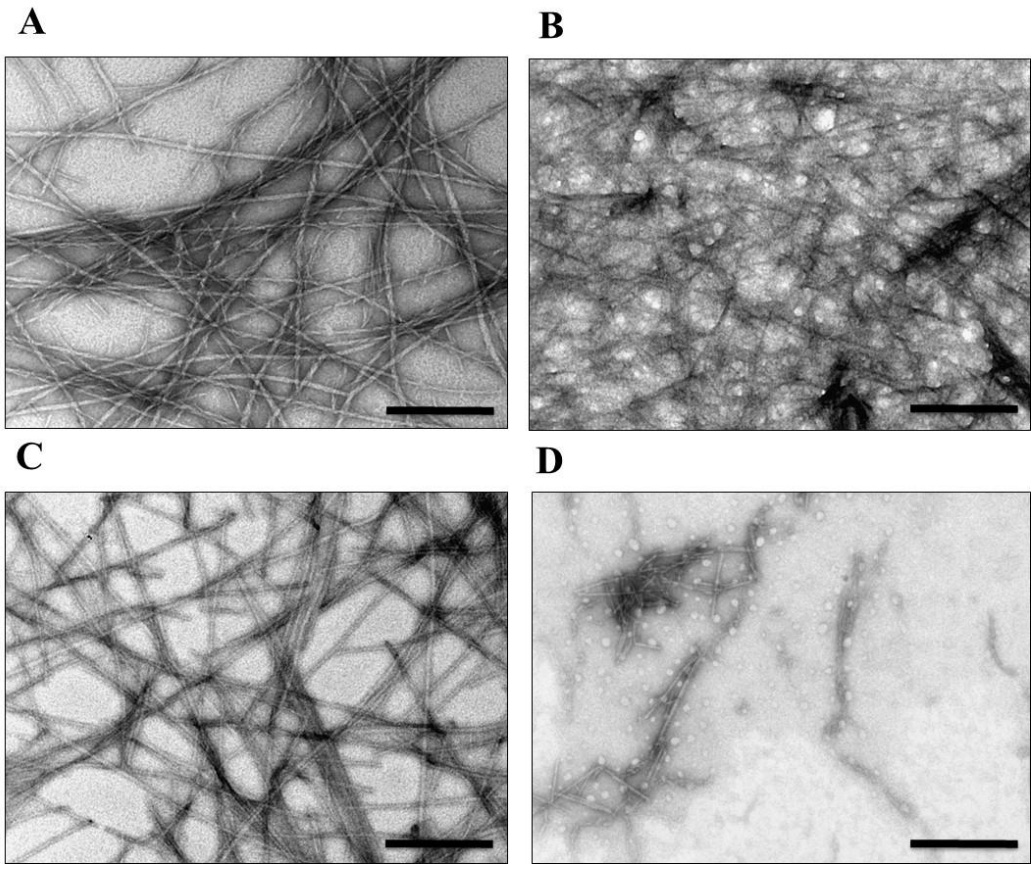


Figure 3.3: *Representative TEM micrographs of peptide acetates in water:(A) Prx3 peptide in water at 40 mg/mL, (B) β-Lac peptide in water at 40 mg/mL, (C) DAPDC peptide in water at 60 mg/mL, (D) UmuD' peptide in water at 100 mg/mL. TEM micrographs correspond to the peptide acetate samples prepared immediately after dilution. (Scale bar is 200 nm)*

For, peptide acetate samples prepared in 100% HFIP or TFA or methanol, no such extended assemblies were observed. These results suggested that the self-assembly of the peptide into extended ordered nanostructures is favored under hydrophilic conditions. It is important to note that the hierarchical structure of the peptide assemblies is stabilized by a large number of noncovalent intermolecular interactions as well as intermolecular peptide-solvent contacts. Also, the major driving force behind folding of proteins and peptides is posited to result from hydrophobic effect or hydrophobic free energy[25]. Earlier de novo designs of self-assembling β-sheet forming peptide sequences highlight that residues such as isoleucine (I), tyrosine (Y),

tryptophan (W), valine (V) and threonine (T) are good β-sheet formers providing the necessary set of hydrophobic interactions alongside the backbone hydrogen bonding[26,27].

Aromatic amino acid residues in short peptide sequences are known to be involved in aromatic stacking through π-π interactions. Short peptide sequences comprised of 5-7 residues containing a single phenylalanine residue self-assembled into amyloid fibrils, similar to those formed by full-length polypeptides[28,29]. Replacement of the phenylalanine residue in the peptide sequence severely impacted the ability of the peptide to self-assemble[30].

Interestingly, all four peptide sequences investigated for self-assembly here have at least one isoleucine residue; valine was present in two out of four peptides, and threonine was present in one out of four peptide sequences. Both β-Lac peptide and the DAPDC peptide contained one phenylalanine residue. The presence of specific residues that promote the β-sheet formation and peptide aggregation explains the intrinsic ability of the designed peptide sequences to self-assemble in solution.

3.5 Characterization of peptide self-assembly in solution

Peptide acetates were soluble in the organic solvent (HFIP), acid (TFA) and polar solvent methanol + water mix and at low concentrations in water. Concentrated peptide acetate solutions at 10 mg/mL for Prx3 peptide and β-Lac peptide, 20 mg/mL for DAPDC peptide and 40 mg/mL for UmuD' peptide prepared in deionized water were found to gel if stored at 4°C over 24 hours. Above 5% w/w (50 mg/mL) for Prx3 peptide and β-Lac peptide, and above 15% w/w (150 mg/mL) for DAPDC peptide and UmuD' peptide, the peptide acetate in water formed highly viscous hydrogels (Figure 3.4).

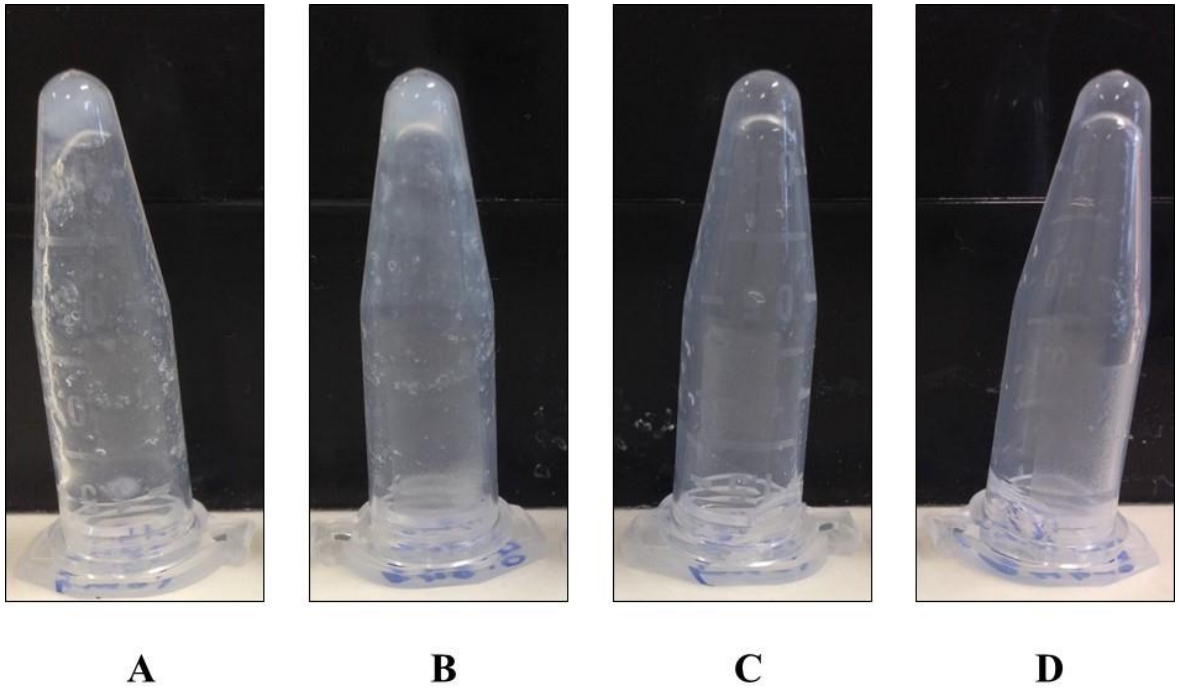


Figure 3.4: *Concentrated peptide acetate salts in water: Prx3 peptide and β-Lac peptide at 60 mg/mL (A and B) and DAPDC peptide & UmuD' peptide at 200 mg/mL (C and D) equilibrated for 24 hours at 4°C. The picture shows viscous hydrogels formed by self-assembly of peptide acetates in water. Prx3 peptide and β-Lac peptide formed highly turbid hydrogels that suggest an extensive network of fibrillar assemblies.*

Gelation is usually associated with water confinement by fibrillar objects that reticulate the media[17]. Turbidity is an indicator of aggregation, more specifically in the case of peptides a representative of the molecular-associated states as observed upon fibrillation in amyloid fibril[31] and in the natural self-assembly of type I collagen[18]. Macroscopic fibrillar assemblies of self-assembling β-sheet type peptides show liquid crystallinity under cross polarizers[32]. Therefore, peptide acetates in water in the 20-200 mg/mL range were examined by light microscopy. All four peptide sequences exhibited birefringence under cross-polarizers with clear liquid crystalline textures (Figure 3.5).

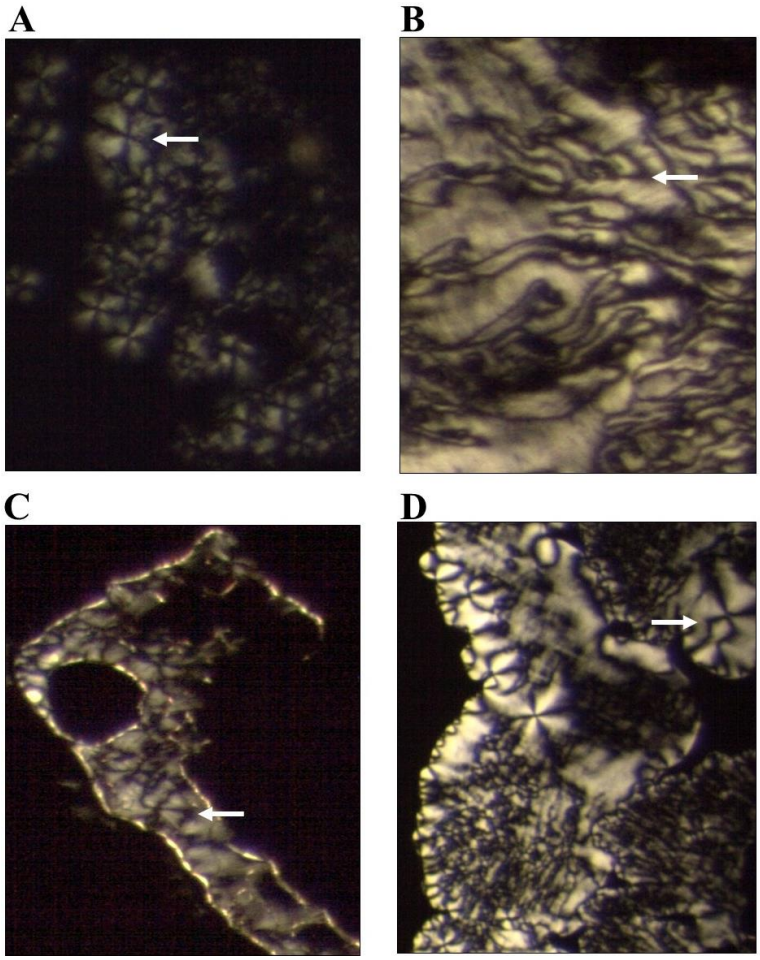


Figure 3.5: *Birefringence of peptide hydrogels observed under polarized light with the cross-polarizers at (± 45°). Optical textures of square columnar phase observed for Prx3 peptide at 20 mg/mL and hexagonal columnar phases observed for the β-Lac peptide (20 mg/mL), DAPDC peptide and UmuD' peptide at 200 mg/mL in water conditioned between two glass slides. The arrows point to one of the highest points of curvature of the specific textures.*

The liquid crystallinity of biological molecules is dependent on their arrangement in a repetitive manner aligned over a long distance, which brings about the elasticity and mechanical strength. Liquid crystals have been earlier reported for reversibly self-assembling peptides[33,34] and this makes them different to amyloid fibrils which are insoluble precipitates. The formation of liquid crystal phases by peptide sequences prone to fibrillate is found to be dependent on the concentration of the peptide, fibril length, pH, and temperature[35,36].

To study the structural organization of the peptide sequences that led to gelation, the peptide acetates in water at dilute concentration were analyzed for secondary structure conformation. Circular dichroism of peptide acetate samples in water equilibrated for over 24 hours exhibited a minimum around 215-220 nm, suggestive of a β-sheet conformation[3] in each case (Figure 3.6).

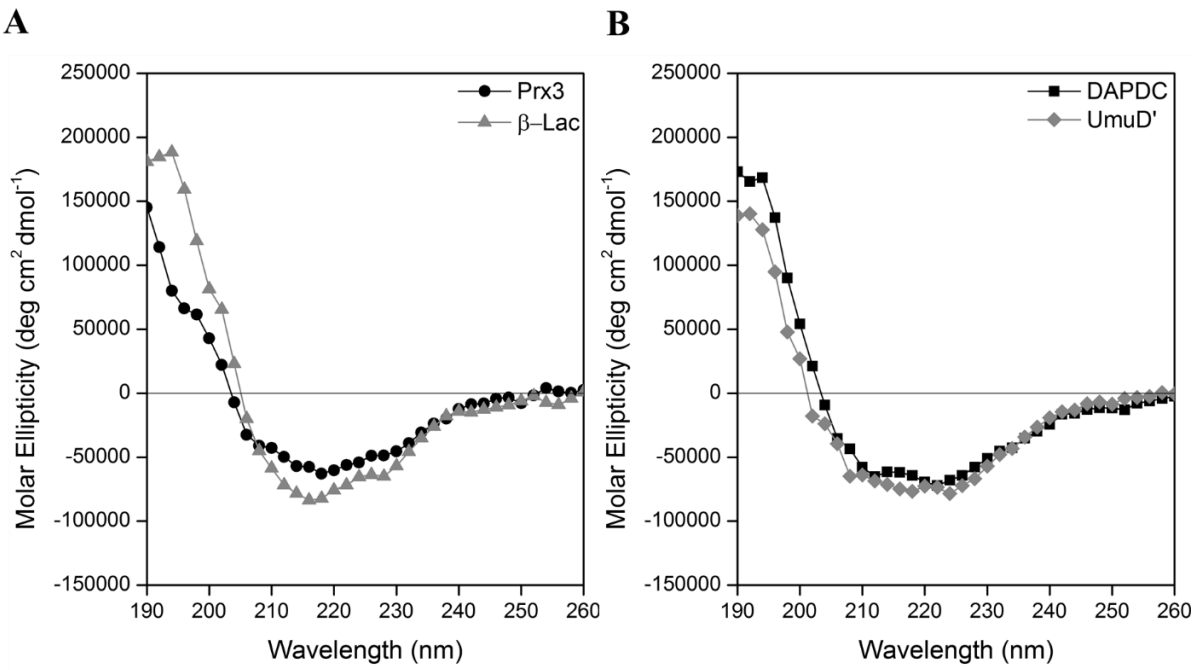


Figure 3.6: *Far UV circular dichroism spectra of peptide acetates in water at 0.008 mg/mL: (A) Prx3 peptide; (B) β-Lac peptide; (C) DAPDC peptide; (D) UmuD' peptide. The peptide acetate in water samples was equilibrated for 24 hours at 4°C and degassed before analysis.*

3.6 ThT binding assays

ThT is a benzothiazole dye known to exhibit enhanced fluorescence intensity in the red region of the visible spectrum (~ 482 nm) upon binding to β-sheet assemblies[38]. Peptide acetates in water in the 20-100 mg/mL range examined immediately after preparation and at 24 hours after equilibration at 4°C, exhibited strong ThT binding (Figure 3.7). This suggested that the macroscopic assembly of the peptides in solution were extended β-sheet based structures.

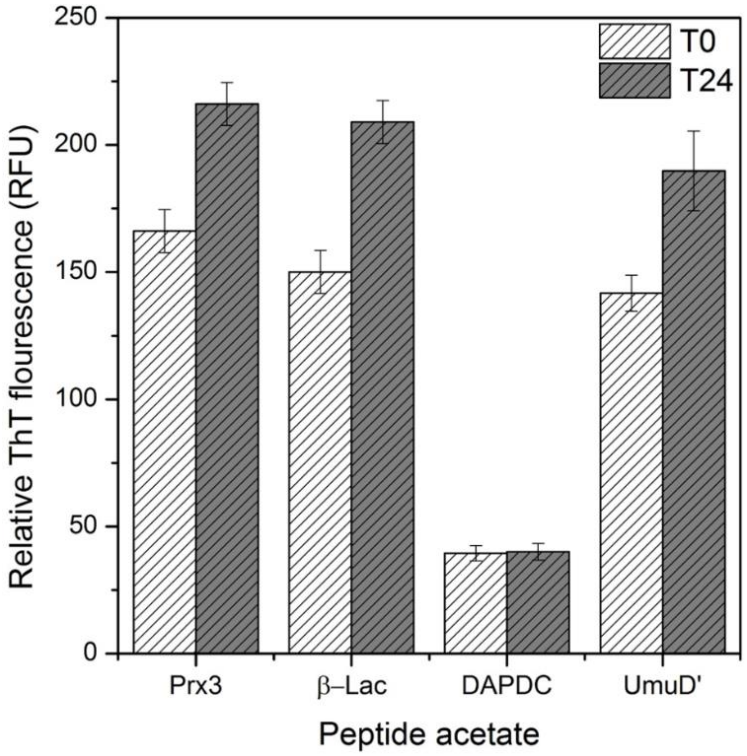


Figure 3.7: *ThT binding assays by dilution for peptide acetates in water at 10 mg/mL recorded immediately after dilution and for samples equilibrated for 24 hours 4°C. The error bars correspond to the standard deviation between triplicate sample preparations.*

ThT binding assays performed by dilution of the peptides from an initially concentrated sample (24 hours old) showed that extended β-sheet structures dissociate upon dilution as indicated by the decrease in ThT fluorescence intensity. ThT binding curves were obtained by plotting the derivative (dt/dc) calculated as the ThT fluorescence intensity per mg of peptide in solution versus peptide acetate concentration (mg mL^{-1}). Two out of the four peptide acetates, the Prx3 peptide, and the β-Lac peptide exhibited a linear derivative curve suggestive of a single state that corresponded to the self-assembled extended β-sheet form of the peptides. The sigmoid growth type derivative curve obtained for the DAPDC peptide and UmuD' peptide, however, was suggestive of a two-state self-assembly process with the presence of a critical concentration of assembly (Figure 3.8).

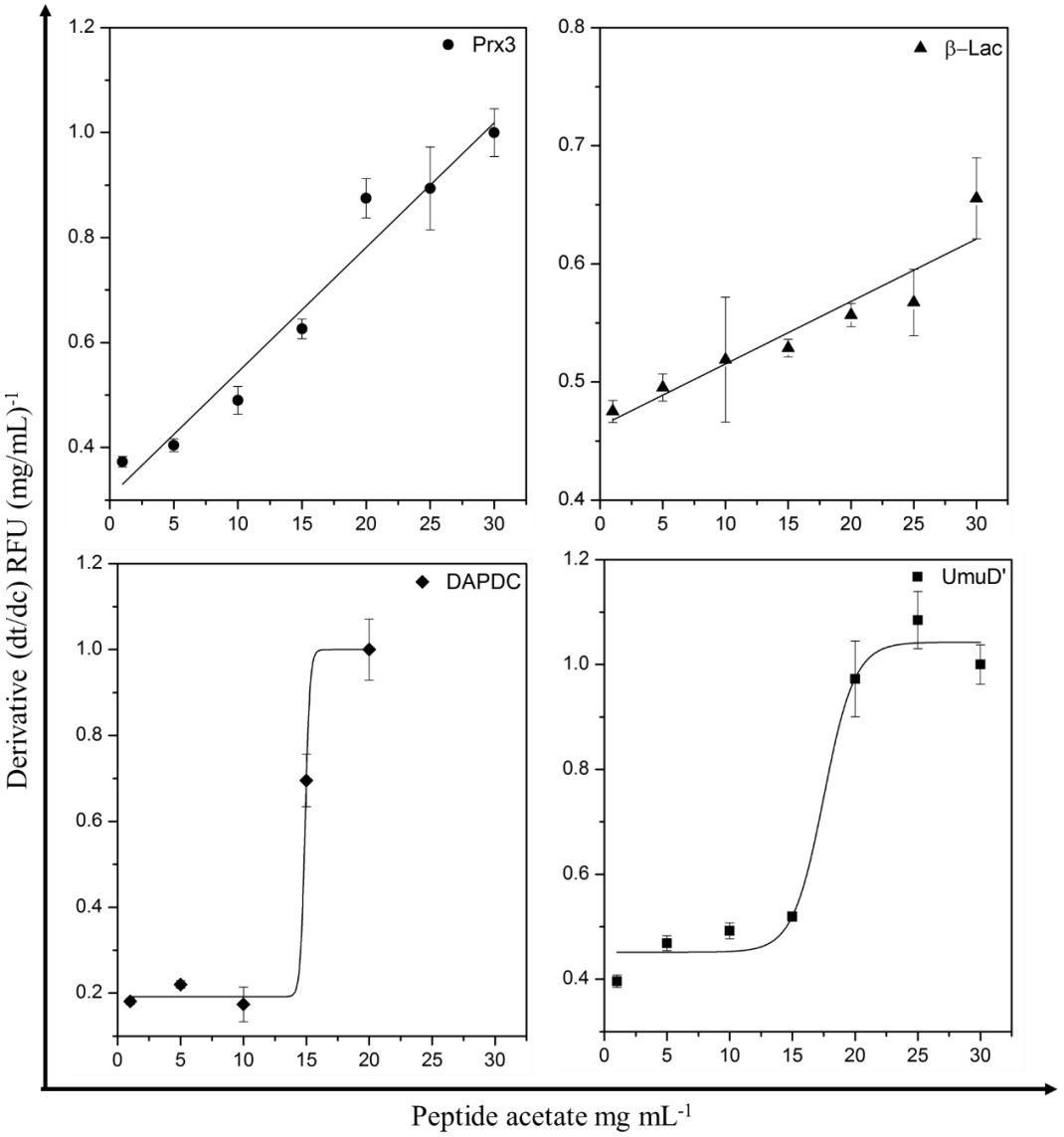


Figure 3.8: *ThT binding assays by dilution for peptide acetates: Fluorescence intensity values recorded 24 hours after dilution for samples equilibrated for 24 hours. The derivative of Prx3 peptide and β-Lac peptide fitted to linear function y = Ax +B. The derivative of DAPDC peptide and UmuD' peptide fitted to sigmoid growth function* $y=B + (A + B)/(1 + e^{(x-x0)/dx})$*. The error bars correspond to the standard deviation between triplicate sample preparations.*

In the case of DAPDC peptide, the peptide acetate in water at > 20 mg/mL, showed large variation in ThT binding measurement, hence was not included in the results. This could be attributed to the increase in macroscopic viscosity of the peptide acetate solution resulting from

the increase in peptide self-assembly into large extended β-sheet structures at high concentration. At concentrations < 15 mg/mL for DAPDC peptide, the steep decline in ThT binding suggested a state of transition between the self-assembled β-sheet state to the smaller prefibrillar aggregates[39].

3.7 TEM characterization of peptide assembly

To study the morphology of the β-sheet rich peptide assemblies in solution, peptide acetates in water in the 40-200 mg/mL range were observed under transmission electron microscope (Section 2.6.8). TEM of negatively stained EM grids of the peptides, revealed that all four peptides self-assembled into nanostructures with extended morphology: nanotapes for Prx3 peptide, bundles of nanofibrils for β-Lac peptide and individual fibrils for DAPDC peptide and UmuD' peptide (Figure 3.9).

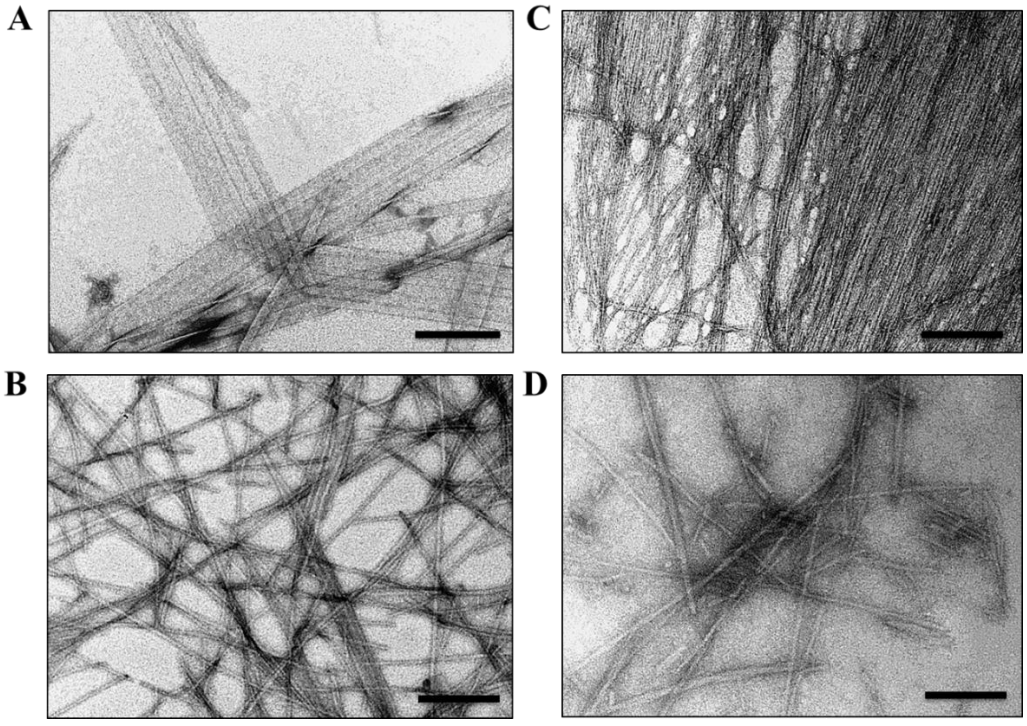


Figure 3.9: *Representative TEM micrographs of (A) Prx3 peptide in water at 40 mg/mL, (B) β-Lac peptide in water at 40 mg/mL, (C) DAPDC peptide in water at 200 mg/mL, (D) UmuD' peptide in water at 100 mg/mL equilibrated for 24 hours. (Scale bar is 200 nm)*

Measurement of nanostructures performed using ImageJ software showed that the extended nanotape like assembly of Prx3 peptide was 8-40 nm wide and of varied lengths between 0.8 to 2 µm (Figure 3.10).

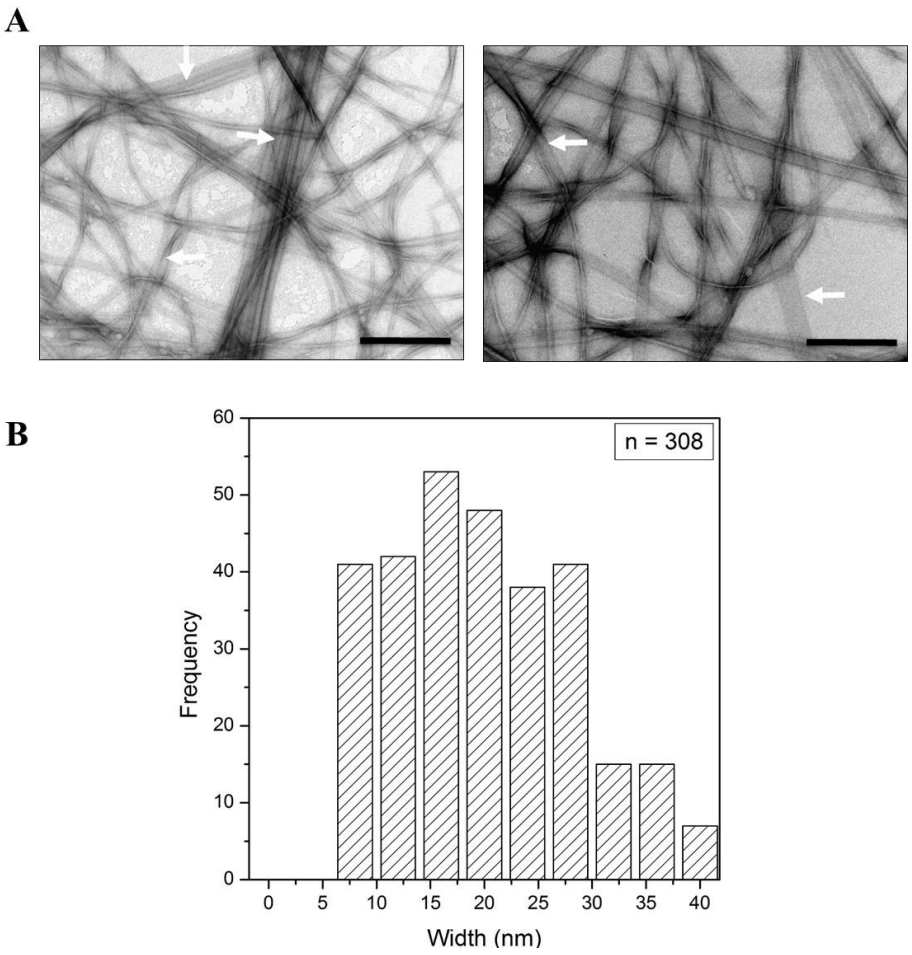


Figure 3.10: *TEM micrographs of Prx3 peptide in water at 40 mg/mL (A). The white arrows point to the different peptide fibril polymorphs. Histogram showing the distribution of fibril width (B) collected from 308 particles respectively. (Scale bar is 100 nm)*

The broad distribution of nanotape width observed suggested that the Prx3 peptide self-assembled into polymorphic nanotapes. Most of the nanostructures spanned across the imaged cross-sectional area of the EM micrograph, which suggested that the observed nanostructures were much larger in length than the estimated values. Polymorphism has been previously described for amyloid fibrils[40,41] and amyloid-like fibrils formed by β-sheet type peptides[42].

Bundles of nanofibrils of β-Lac peptide were on an average 2 μm in length. (Figure 3.9 B). DAPDC peptide fibrils were between 0.6-1.5 μm in length (Figure 3.9 C) however, fibrils of UmuD' peptide were much smaller, with a length distribution between 0.2-1.0 μm. Fibrils with a length larger than 1 μm were uncommon with only a few seen spanning across the imaged micrograph cross-sectional area (Figure 3.9 D). The peptide fibrils for both DAPDC peptide and UmuD' peptide were 4-10 nm in width. The distribution of width observed in the case of the peptides can be explained by the intertwining of fibrils that form a thick bundle of fibrils.

3.8 Characterization of peptide self-assembly at alkaline pH

Amino acid residues such as histidine contain ionizable imidazole side chain group with a pK_a of 6.0. To investigate the effect of a change in the protonated state of the histidine side chain in the self-assembly of the peptides, they were studied at alkaline pH. CD spectra of peptide acetates dissolved in imidazole buffer at pH 8.5 were found to be similar to those observed for peptide acetates in water at ~ pH 5.0 with the characteristic minimum around 215-230 nm (Figure 3.11).

The circular dichroism spectra of the peptide acetates observed were in agreement with the different studies published on peptide sequences that self-assemble into β-sheets. Also, theoretical studies on the optical properties of α-helices and β-sheets observed in solution have shown that the negative minimum that is consistent with an n to π* transition is observed for β-sheets in the 215-230 nm wavelength range[43]. The result thus suggested that change in the protonated state of histidine as a function of pH does not have a substantial effect on the secondary structure conformation of the peptide sequences in solution.

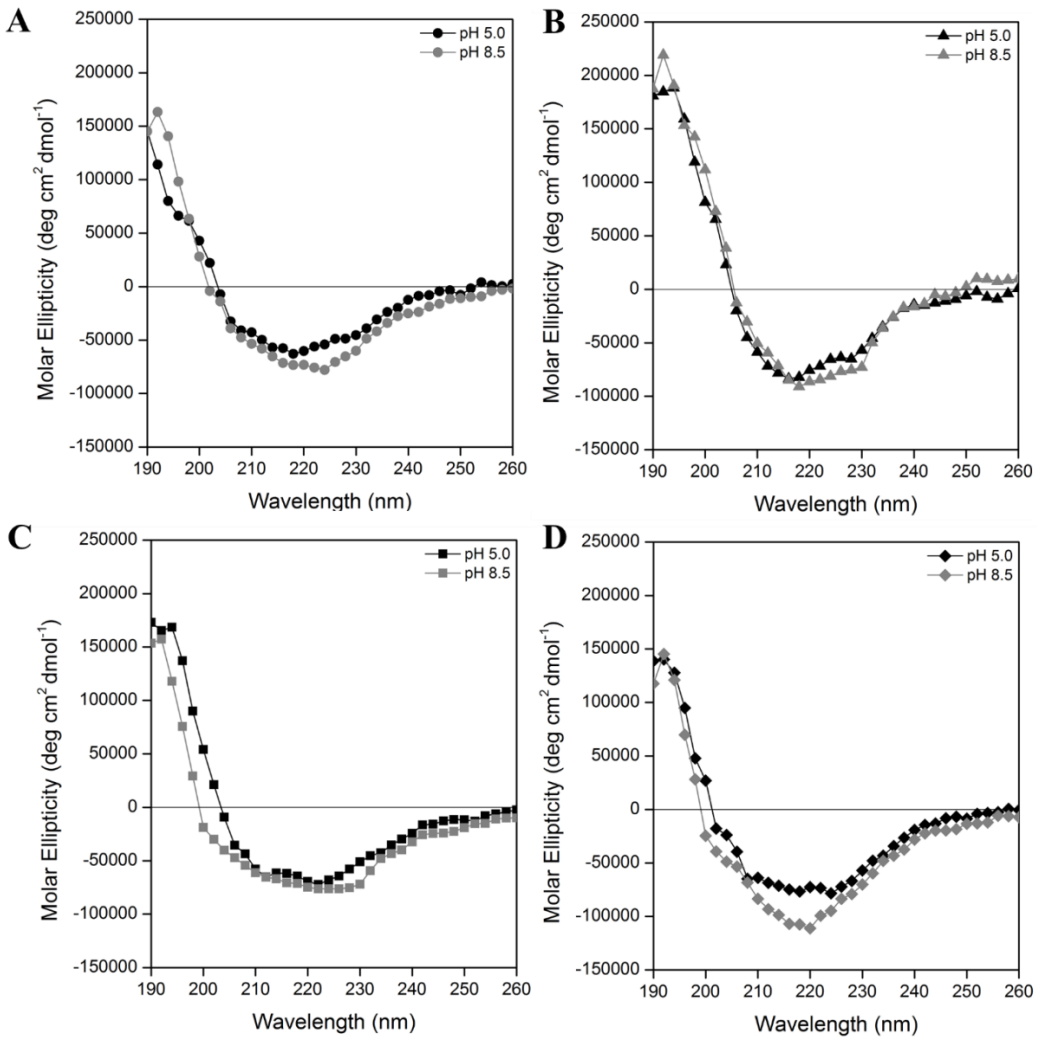


Figure 3.11: *Overlaid CD spectra of peptide acetates :(A) Prx3 peptide; (B) β-Lac peptide; (C) DAPDC peptide; (D) UmuD' peptide in water (pH 5.0) and in 100 mM imidazole (pH 8.5) at 0.008 mg/mL. The peptide acetate samples were equilibrated for 24 hours at 4°C and degassed before analysis.*

TEM analysis of peptide acetates in imidazole pH 8.5 in the 20-100 mg/mL range revealed self-assembled peptide nanostructures similar to those observed for the peptides in water (Figure 3.12 and 3.13 panel 1). Dilute samples of the peptide acetate in imidazole at 1.0 mg/mL for Prx3 peptide and β-Lac peptide, 10 mg/mL for DAPDC peptide and 20 mg/mL for UmuD' peptide revealed nanostructures of much smaller dimensions, consistent with the peptide acetate in water samples (Figure 3.12 and 3.13 panel 2).

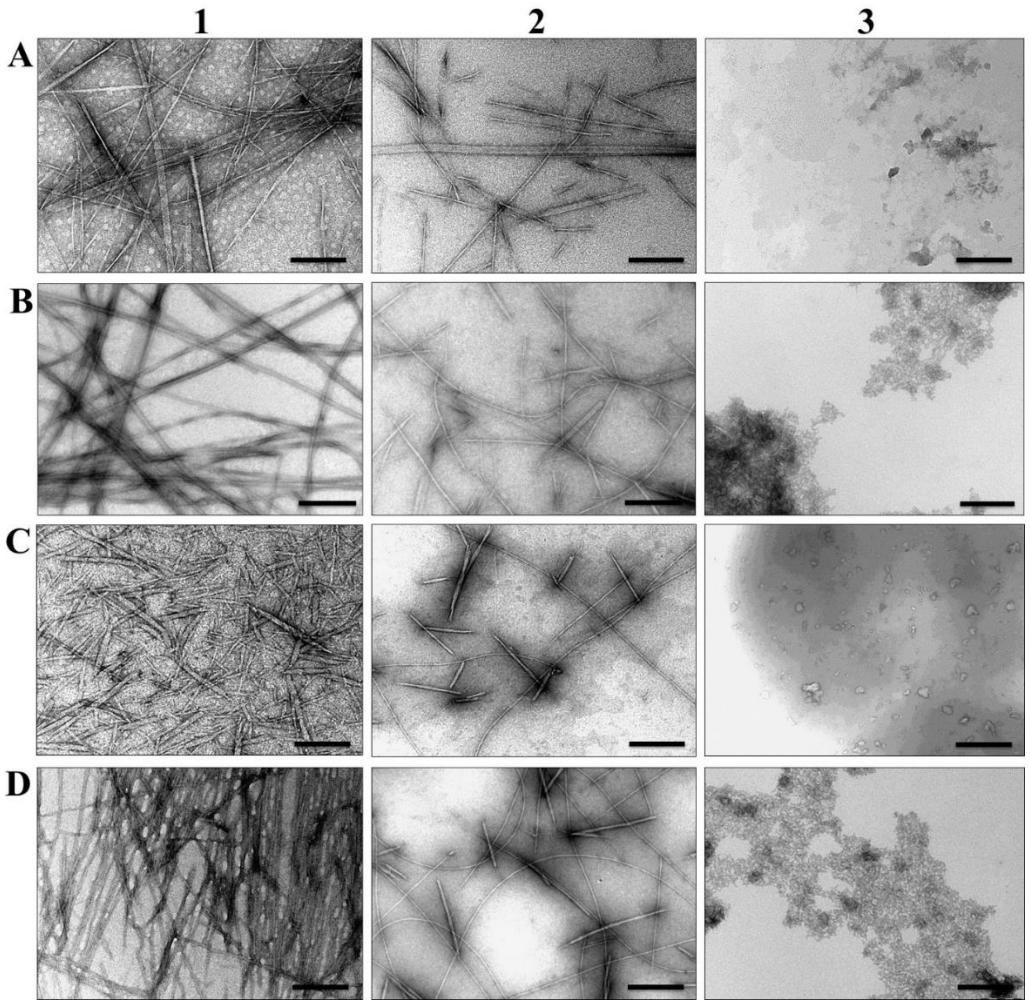


Figure 3.12: *TEM micrographs of peptide acetates Prx3 and β-Lac in 100 mM imidazole pH 8.5 (A,C) and in water (B,D) equilibrated for 24 hours. Numbers correspond to peptide concentration at 20 mg/mL (1), 1 mg/mL (2) and 0.01 mg/mL (3). (Scale bar is 200 nm)*

At the most diluted state; 0.01 mg/mL for Prx3 peptide and β-Lac peptide, 1 mg/mL for DAPDC peptide and UmuD' peptide in imidazole buffer, no such extended nanostructures were observed (Figure 3.12 and 3.13 panel 3). This suggested the dissociation of extended nanostructures upon dilution, consistent with ThT binding assays.

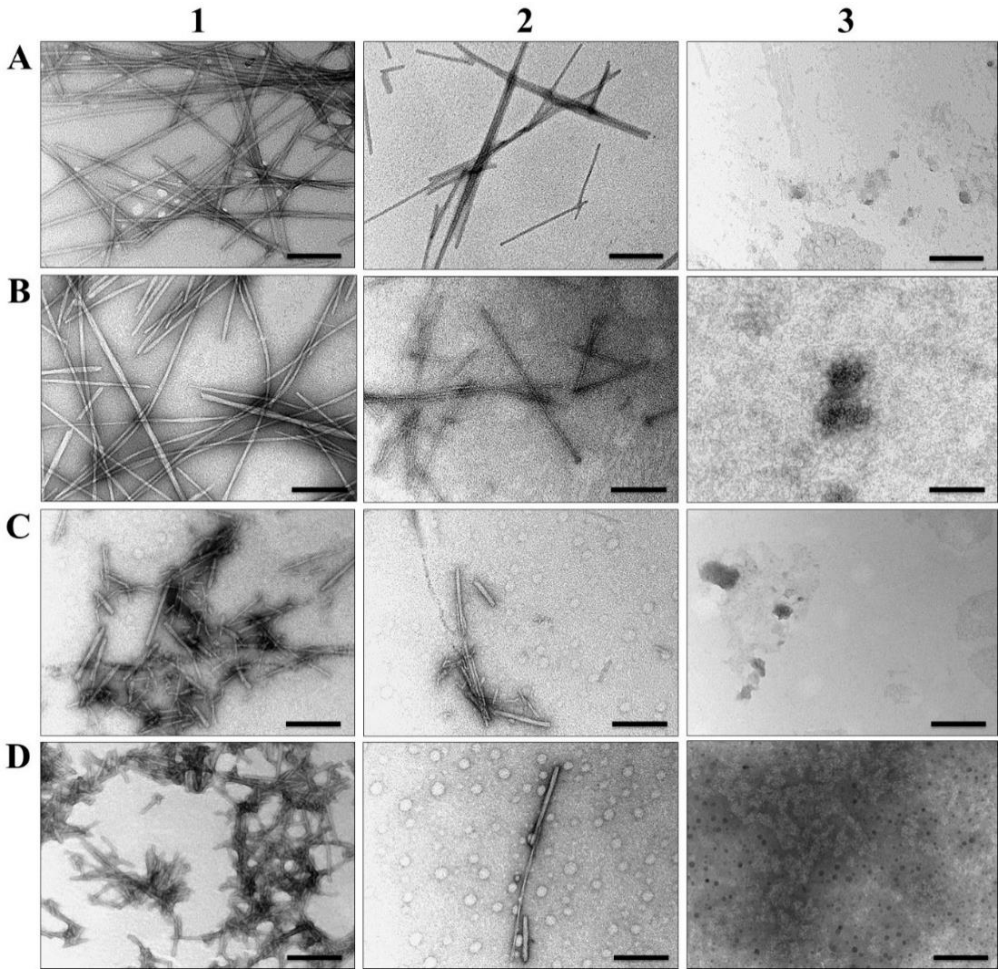


Figure 3.13: *TEM micrographs of peptide acetates DAPDC and UmuD' in 100 mM imidazole pH 8.5 (A,C) and in water (B,D) equilibrated for 24 hours. Numbers correspond to peptide concentrations at 50 mg/mL (1), 10 mg/mL (2), 1 mg/mL (3) for DAPDC peptide and 100 mg/mL (1), 20 mg/mL (2) and 1 mg/mL (3) for UmuD' peptide. (Scale bar is 200 nm).*

3.9 Small angle X-ray scattering of peptides

The molecular architecture within the extended nanostructures formed by peptides in solution was examined by SAXS. Concentrated sample of the peptide acetates in water in the 100-200 mg/mL range was analyzed at the SAXS beamline at the Australian Synchrotron Facility (Section 2.6.9). The scattering profile obtained by plotting the log of the SAXS intensity (A.U.) versus the q value ($Å^{-1}$) exhibited a distinct peak at 1.3 $Å^{-1}$ consistent for all the peptides (Figure 3.14).

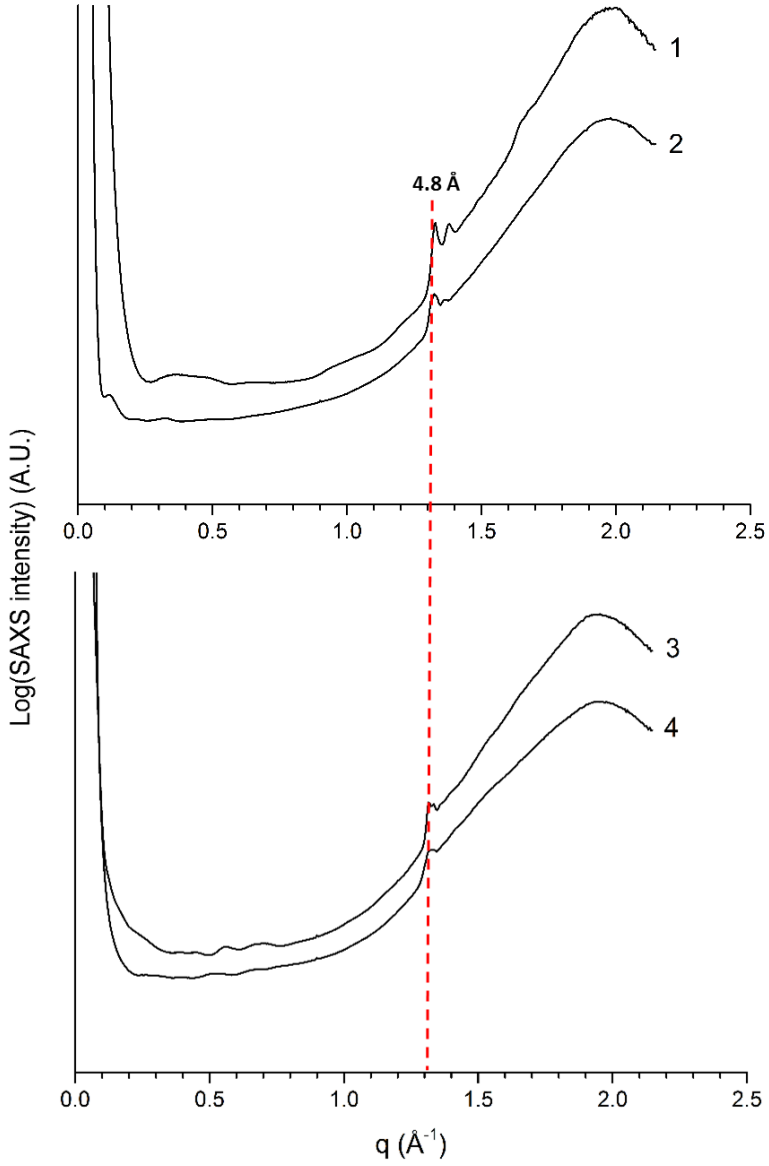


Figure 3.14: *SAXS patterns recorded at room temperature for peptide acetate in water: (1) Prx3 peptide; (2) β-peptide; (3) DAPDC peptide; (4) UmuD' peptide. The dashed red line shows the reflection at 1.3 Å^{-1} visible for all the peptide sequences.*

The presence of Braggs reflection at q = 1.3 Å^{-1} equates to the real space distance of ~ 4.8 Å (d = 2π/q), which corresponds to the characteristic β-sheet intra-strand distance[44]. This suggested that the peptide self-assembled in solution to form extended β-sheet based assemblies. The high intensity of the peak indicated that this distance corresponded to the repetitive distance in the extended intermolecular network of β-sheets.

X-ray scattering studies on short peptides designed from motifs within proteins associated with amyloidosis, show characteristics meridional reflections at q value = 1.3 $Å^{-1}$ and correspond to the hydrogen bonding distance in the β-sheet[45,46]. The lesser defined reflections observed at low angles > 0.5 $Å^{-1}$ was attributed to the diffused scattering of the samples. Such SAXS patterns have been previously described in the case of the fibril forming neuropeptide somatostatin-14[47] and are thought to arise from the anisotropic form factor of fibrillar nanostructures perpendicular to the fibril axis[48]. Overall, the SAXS data showed that the peptide sequences self-assembled in solution into extended β-sheet based nanostructures that are similar in their molecular organization.

3.10 Discussion

The use of peptide tectons to produce molecular biomaterials with specific properties that allow control over self-assembly is central to bionanotechnology. Different self-assembling peptide sequences designed to date have largely focused on a limited set of sequences of biological origin. This research work aimed to introduce a new design source of peptide sequences that self-assemble in solution.

Peptide sequences, were designed from simple protein-β interfaces of four different classes of protein homo-oligomers and investigated for self-assembly. All the peptide sequences in water formed birefringent hydrogels, with distinct liquid crystalline textures. The gelation of peptide acetate solutions was consistent with higher order assembly of the peptides. Birefringence observed for the peptide hydrogels was suggestive of the nematic phase of the liquid crystalline assemblies. The nematic liquid crystal phase is characteristic of molecules with a definitive order or pattern[49,50]. Electron microscopy revealed that all peptide sequences self-assembled spontaneously in solution into nanostructures with elongated morphologies. Such elongated nanostructures fall within the range of different morphologies described previously for β-sheet type self-assembling peptides[51].

EM further showed that the extended nanostructures formed by Prx3 peptide exhibited polymorphism. Polymorphism is reported to be an inherent property of amyloid fibrils [40,41,52], with such differences thought to arise due to the different ways of association and packing mode of building block units[51]. In a recently published work, MAS NMR coupled with cryo-EM reconstruction showed that the polymorphic fibrils formed by peptide sequences derived from transthyretin protein comprised of β-sheets organized as two, three or four sheets laterally associated through terminal side-chain interactions in the extended nanoribbon architecture[42]. This can explain the type of distribution of nanotape width measured for the Prx3 peptide.

The intertwining of peptide nanoribbons and nanofibrils observed in the case of Prx3 peptide and DAPDC peptide could be attributed to the twisting of the protofilament to minimize exposure of hydrophobic patches, as earlier described for fibrillization of amyloid Aβ peptide[53], α-synuclein[54] and insulin[55]. The periodicity of intertwining observed in the extended peptide nanostructures can be explained by the balance between the mechanical forces dominated by the elasticity of the protofilament and the electrostatic interactions arising from the distribution of hydrophobic and charge residues[53].

The four peptide sequences in water under dilute concentration exhibited the expected propensity for β-sheet secondary structure conformation[37]. For concentrated peptide hydrogels, the distinct Braggs reflection centered on q value = 1.3 $Å^{-1}$ consistent in all peptide acetates corresponded to extended β-sheet networks arranged in a repetitive manner over long distances[46,56,57].

The extended β-sheet networks in all the four cases were found to dissociate upon dilution as suggested by the ThT binding assays for dilute peptide acetate in water solutions. Derivative of ThT fluorescence intensity to peptide acetate was used to predict the self-assembly models ThT derivative Furthermore, absence of a linear distribution parallel to the x-axis for the ThT derivative plots suggested that the decrease in ThT fluorescence intensity is not an effect of dilution but due to the dissociation of the peptide β-sheet structures. Reversibility of assembly has been previously described for some peptide sequences[47,58,59]. In the case of UmuD' peptide, the β-sheet assemblies estimated by ThT binding assays were found to be much smaller in length compared to the other three peptides as shown by TEM. The ThT assay further suggested a higher critical assembly concentration for the DAPDC peptide compared to the other peptides.

Interestingly, the ThT results correlated well with the type of oligomer interface from which the peptides were designed. The Prx3 peptide, the β-Lac peptide and the UmuD' peptide correspond to the homodimer interface of the protein, with the K_d of the homodimer interface reported to be in the micromolar range for both Prx3[60] and β-Lac[61], and in picomolar range for *E. coli* UmuD'protein[62]. It is important to note that in the high-resolution NMR structure of the UmuD' homodimer, hydrophobic pockets can be seen present near the protein homodimer β-interface. The hydrophobic interactions along with steric effects arising from neighboring residues along the interface favor the formation of the protein homodimer at a much lower concentration in solution[11]. In the isolated form (UmuD' peptide), however the peptide sequence is not able to reproduce the full properties of the interface motif, which explains the relatively higher critical assembly concentration observed for the UmuD' peptide.

The DAPDC peptide corresponds to the homotetramer interface of *M. tuberculosis* DAPDC protein that is still under debate, though posited to exist at very high concentrations within the protein crystal[10]. The, K_d of the interface is still disputed but thought to be in the micromolar range[12], which explains the higher critical assembly concentration observed for the DAPDC peptide. The results shown here supports the connection that the derived peptide sequences are tectons with intrinsic assembly properties derived from the native parent interfaces.

This work introduces a new design source for peptide tectons and adds to the current library of self-assembling β-sheet type peptide sequences. Such peptides inspired from naturally optimized protein interfaces can prove to be excellent candidates as building blocks for synthetic self-assembly.

3.11 References

1. Rajagopal, K. & Schneider, J. P. (2004). Self-assembling peptides and proteins for nanotechnological applications. *Current Opinion in Structural Biology* **14**, 480-486.
2. Matthews, J. M. (2012). Protein dimerization and oligomerization in biology. *Springer Science*.
3. Dey, S., Pal, A., Chakrabarti, P. & Janin, J. (2010). The subunit interfaces of weakly associated homodimeric proteins. *Journal of Molecular Biology* **398**, 146-160.
4. Nooren, I. M. & Thornton, J. M. (2003). Diversity of protein-protein interactions. *EMBO Journal* **22**, 3486-3492.
5. Bromley, E. H., Channon, K., Moutevelis, E. & Woolfson, D. N. (2008). Peptide and protein building blocks for synthetic biology: from programming biomolecules to self-organized biomolecular systems. *ACS Chemical Biology* **3**, 38-50.
6. Guharoy, M. & Chakrabarti, P. (2007). Secondary structure based analysis and classification of biological interfaces: identification of binding motifs in protein-protein interactions. *Bioinformatics* **23**, 1909-1918.
7. Cao, Z., Roszak, A. W., Gourlay, L. J., Lindsay, J. G. & Isaacs, N. W. (2005). Bovine mitochondrial peroxiredoxin III forms a two-ring catenane. *Structure* **13**, 1661-1664.
8. Adams, J. J., Anderson, B. F., Norris, G. E., Creamer, L. K. & Jameson, G. B. (2006). Structure of bovine β-lactoglobulin (variant A) at very low ionic strength. *Journal of Structural Biology* **154**, 246-254.
9. Vijayalakshmi, L., Krishna, R., Sankaranarayanan, R. & Vijayan, M. (2008). An asymmetric dimer of β-lactoglobulin in a low humidity crystal form-structural changes that accompany partial dehydration and protein action. *Proteins: Structure, Function, and Bioinformatics* **71**, 241-249.
10. Kefala, G., Perry, L. J. & Weiss, M. S. (2005). Cloning, expression, purification, crystallization and preliminary X-ray diffraction analysis of LysA (Rv1293) from *Mycobacterium tuberculosis*. *Acta Crystallographica Section F: Structural Biology and Crystallization Communications* **61**, 782-784.
11. Ferentz, A. E., Walker, G. C. & Wagner, G. (2001). Converting a DNA damage checkpoint effector (UmuD2C) into a lesion bypass polymerase (UmuD'2C). *EMBO Journal* **20**, 4287-4298.

12. Weyand, S., Kefala, G., Svergun, D. I. & Weiss, M. S. (2009). The three-dimensional structure of diaminopimelate decarboxylase from Mycobacterium tuberculosis reveals a tetrameric enzyme organisation. *Journal of Structural and Functional Genomics* **10**, 209-217.
13. Hamada, D. & Dobson, C. M. (2002). A kinetic study of β-lactoglobulin amyloid fibril formation promoted by urea. *Protein Science* **11**, 2417-2426.
14. Bowerman, C. J. & Nilsson, B. L. (2012). Review self-assembly of amphipathic β-sheet peptides: insights and applications. *Peptide Science* **98**, 169-184.
15. Morris, Kyle L., Rodger, A., Hicks, Matthew R., Debulpaep, M., Schymkowitz, J., Rousseau, F. & Serpell, Louise C. (2013). Exploring the sequence–structure relationship for amyloid peptides. *Biochemical Journal* **450**, 275-283.
16. Lee, N. R., Bowerman, C. J. & Nilsson, B. L. (2013). Sequence length determinants for self-assembly of amphipathic β-sheet peptides. *Biopolymers* **100**, 738-750.
17. de Groot, N. S., Parella, T., Aviles, F. X., Vendrell, J. & Ventura, S. (2007). Ile-Phe dipeptide self-assembly: clues to amyloid formation. *Biophysical Journal* **92**, 1732-1741.
18. Kar, K., Wang, Y.-H. & Brodsky, B. (2008). Sequence dependence of kinetics and morphology of collagen model peptide self-assembly into higher order structures. *Protein Science* **17**, 1086-1095.
19. Maezawa, I., Hong, H.-S., Liu, R., Wu, C.-Y., Cheng, R. H., Kung, M.-P., Kung, H. F., Lam, K. S., Oddo, S., LaFerla, F. M. & Jin, L.-W. (2008). Congo red and thioflavin-T analogs detect Aβ oligomers. *Journal of Neurochemistry* **104**, 457-468.
20. LeVine, H., 3rd. (1993). Thioflavin-T interaction with synthetic Alzheimer's disease β-amyloid peptides: detection of amyloid aggregation in solution. *Protein Science* **2**, 404-10.
21. LeVine, H., 3rd. (1999). Quantification of β-sheet amyloid fibril structures with thioflavin T. *Methods in Enzymology* **309**, 274-84.
22. Gras, S. L., Waddington, L. J. & Goldie, K. N. (2011). Transmission electron microscopy of amyloid fibrils. *Methods in Molecular Biology* **752**, 197-214.
23. Sabatini, D. D., Bensch, K. & Barrnett, R. J. (1963). Cytochemistry and electron microscopy. The preservation of cellular ultrastructure and enzymatic activity by aldehyde fixation. *Journal of Cell Biology* **17**, 19-58.

24. Pras, M., Schubert, M., Zucker-Franklin, D., Rimon, A. & Franklin, E. C. (1968). The characterization of soluble amyloid prepared in water. *Journal of Clinical Investigation* **47**, 924-933.
25. Kauzmann, W. (1959). Some factors in the interpretation of protein denaturation. In advances in protein chemistry. *Academic Press* **14**, 1-63.
26. Chou, P. Y. & Fasman, G. D. (1974). Conformational parameters for amino acids in helical, β-sheet, and random coil regions calculated from proteins. *Biochemistry* **13**, 211-222.
27. Kim, C. A. & Berg, J. M. (1993). Thermodynamic β-sheet propensities measured using a zinc-finger host peptide. *Nature* **362**, 267-270.
28. Tenidis, K., Waldner, M., Bernhagen, J., Fischle, W., Bergmann, M., Weber, M., Merkle, M.-L., Voelter, W., Brunner, H. & Kapurniotu, A. (2000). Identification of a penta- and hexapeptide of islet amyloid polypeptide (IAPP) with amyloidogenic and cytotoxic properties1. *Journal of Molecular Biology* **295**, 1055-1071.
29. Reches, M., Porat, Y. & Gazit, E. (2002). Amyloid fibril formation by pentapeptide and tetrapeptide fragments of human calcitonin. *Journal of Biological Chemistry* **277**, 35475-35480.
30. Gazit, E. (2007). Self-assembly of short aromatic peptides into amyloid fibrils and related nanostructures. *Prion* **1**, 32-35.
31. Cook, N. P. & Martí, A. A. (2012). Facile methodology for monitoring amyloid-β fibrillization. *ACS Chemical Neuroscience* **3**, 896-899.
32. Valéry, C., Paternostre, M., Robert, B., Gulik-Krzywicki, T., Narayanan, T., Dedieu, J.-C., Keller, G., Torres, M.-L., Cherif-Cheikh, R., Calvo, P. & Artzner, F. (2003). Biomimetic organization: octapeptide self-assembly into nanotubes of viral capsid-like dimension. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 10258-10262.
33. Valery, C., Artzner, F., Robert, B., Gulick, T., Keller, G., Grabielle-Madelmont, C., Torres, M. L., Cherif-Cheikh, R. & Paternostre, M. (2004). Self-association process of a peptide in solution: from β-sheet filaments to large embedded nanotubes. *Biophysical Journal* **86**, 2484-2501.
34. Aggeli, A., Bell, M., Carrick, L. M., Fishwick, C. W., Harding, R., Mawer, P. J., Radford, S. E., Strong, A. E. & Boden, N. (2003). pH as a trigger of peptide β-sheet self-assembly and reversible switching between nematic and isotropic phases. *Journal of the American Chemical Society* **125**, 9619-9628.

35. Fichman, G. & Gazit, E. (2014). Self-assembly of short peptides to form hydrogels: design of building blocks, physical properties and technological applications. *Acta Biomaterialia* **10**, 1671-1682.

36. Pomerantz, W. C., Yuwono, V. M., Pizzey, C. L., Hartgerink, J. D., Abbott, N. L. & Gellman, S. H. (2008). Nanofibers and lyotropic liquid crystals from a class of self-assembling β-peptides. *Angewandte Chemie Intertional Edition in English* **47**, 1241-1244.

37. Greenfield, N. J. (2006). Using circular dichroism spectra to estimate protein secondary structure. *Nature protocols* **1**, 2876-2890.

38. Sulatskaya, A. I., Kuznetsova, I. M. & Turoverov, K. K. (2012). Interaction of Thioflavin T with amyloid fibrils: fluorescence quantum yield of bound dye. *Journal of Physical Chemistry B* **116**, 2538-2544.

39. Otzen, D. E. (2013). Amyloid fibrils and prefibrillar aggregates: molecular and biological properties, *John Wiley & Sons*.

40. Goldsbury, C. S., Cooper, G. J., Goldie, K. N., Muller, S. A., Saafi, E. L., Gruijters, W. T., Misur, M. P., Engel, A., Aebi, U. & Kistler, J. (1997). Polymorphic fibrillar assembly of human amylin. *Journal of Structural Biology* **119**, 17-27.

41. Petkova, A. T., Leapman, R. D., Guo, Z., Yau, W. M., Mattson, M. P. & Tycko, R. (2005). Self-propagating, molecular-level polymorphism in Alzheimer's β-amyloid fibrils. *Science* **307**, 262-265.

42. Fitzpatrick, A. W. P., Debelouchina, G. T., Bayro, M. J., Clare, D. K., Caporini, M. A., Bajaj, V. S., Jaroniec, C. P., Wang, L., Ladizhansky, V., Müller, S. A., MacPhee, C. E., Waudby, C. A., Mott, H. R., De Simone, A., Knowles, T. P. J., Saibil, H. R., Vendruscolo, M., Orlova, E. V., Griffin, R. G. & Dobson, C. M. (2013). Atomic structure and hierarchical assembly of a cross-β amyloid fibril. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 5468-5473.

43. Woody, R. W. & Koslowski, A. (2002). Recent developments in the electronic spectroscopy of amides and α-helical polypeptides. *Biophysical Chemistry* **102**, 535-551.

44. Segman, S., Lee, M.-r., Vaiser, V., Gellman, S. H. & Rapaport, H. (2010). Highly stable pleated-sheet secondary structure in assemblies of amphiphilic α/β-peptides at the air–water Interface. *Angewandte Chemie International Edition* in English **49**, 716-719.

45. Langkilde, A. E. & Vestergaard, B. (2009). Methods for structural characterization of prefibrillar intermediates and amyloid fibrils. *FEBS Letters* **583**, 2600-2609.

46. Langkilde, A. E., Morris, K. L., Serpell, L. C., Svergun, D. I. & Vestergaard, B. (2015). The architecture of amyloid-like peptide fibrils revealed by X-ray scattering, diffraction and electron microscopy. *Acta Crystallographica Section D: Biological Crystallography* **71**, 882-895.

47. van Grondelle, W., Iglesias, C. L., Coll, E., Artzner, F., Paternostre, M., Lacombe, F., Cardus, M., Martinez, G., Montes, M., Cherif-Cheikh, R. & Valery, C. (2007). Spontaneous fibrillation of the native neuropeptide hormone somatostatin-14. *Journal of Structural Biology* **160**, 211-223.

48. Guinier, A. (1994). X-ray diffraction in crystals, imperfect crystals, and amorphous bodies, *Dover*.

49. Janek, K., Behlke, J., Zipper, J., Fabian, H., Georgalis, Y., Beyermann, M., Bienert, M. & Krause, E. (1999). Water-soluble β-sheet models which self-assemble into fibrillar structures. *Biochemistry* **38**, 8246-8252.

50. Jin, L.-W., Claborn, K. A., Kurimoto, M., Geday, M. A., Maezawa, I., Sohraby, F., Estrada, M., Kaminksy, W. & Kahr, B. (2003). Imaging linear birefringence and dichroism in cerebral amyloid pathologies. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 15294-15298.

51. Castillo, J., Sasso, L. & Svendsen, W. E. (2012). Self-assembled peptide nanostructures: advances and applications in nanobiotechnology. *Pan Stanford*.

52. Pedersen, J. S., Andersen, C. B. & Otzen, D. E. (2010). Amyloid structure-one but not the same: the many levels of fibrillar polymorphism. *FEBS Journal* **277**, 4591-4601.

53. van Gestel, J. & de Leeuw, S. W. (2007). The eormation of fibrils by intertwining of filaments: model and application to amyloid Aβ protein. *Biophysical Journal* **92**, 1157-1163.

54. van Raaij, M. E., van Gestel, J., Segers-Nolten, I. M. J., de Leeuw, S. W. & Subramaniam, V. (2008). Concentration dependence of α-synuclein fibril Length assessed by quantitative atomic force microscopy and statistical-mechanical theory. *Biophysical Journal* **95**, 4871-4878.

55. Jiménez, J. L., Nettleton, E. J., Bouchard, M., Robinson, C. V., Dobson, C. M. & Saibil, H. R. (2002). The protofilament structure of insulin amyloid fibrils. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 9196-9201.

56. Sunde, M., Serpell, L. C., Bartlam, M., Fraser, P. E., Pepys, M. B. & Blake, C. C. F. (1997). Common core structure of amyloid fibrils by synchrotron X-ray diffraction. *Journal of Molecular Biology* **273**, 729-739.

57. Serpell, L. C. (2000). Alzheimer's amyloid fibrils: structure and assembly. *Biochimica et Biophysica Acta* **1502**, 16-30.

58. Valery, C., Artzner, F. & Paternostre, M. (2011). Peptide nanotubes: molecular organisations, self-assembly mechanisms and applications. *Soft Matter* **7**, 9583-9594.

59. Andersen, K. B., Castillo-Leon, J., Hedstrom, M. & Svendsen, W. E. (2011). Stability of diphenylalanine peptide nanotubes in solution. *Nanoscale* **3**, 994-998.

60. Barranco-Medina, S., Lazaro, J. J. & Dietz, K. J. (2009). The oligomeric conformation of peroxiredoxins links redox state to function. *FEBS Letters* **583**, 1809-16.

61. Owusu Apenten, R. K. & Galani, D. (2000). Thermodynamic parameters for β-lactoglobulin dissociation over a broad temperature range at pH 2.6 and 7.0. *Thermochimica Acta* **359**, 181-188.

62. Simon, S. M., Sousa, F. J. R., Mohana-Borges, R. & Walker, G. C. (2008). Regulation of *E. coli* SOS mutagenesis by dimeric intrinsically disordered UmuD gene products. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 1152-1157.

4 Chapter Four

Alteration of Prx3 protein-β interface peptide by sequence modification

4.1 Introduction

Self-assembling peptides have shown much promise as molecular tectons for different applications in bionanotechnology such as metal templating, biomineralization, biosensing, drug amphiphile for targeted drug delivery and for fabrication of nanodevices. For many of these applications, tuning, and control of the self-assembly process is highly desirable. Subtle sequence changes such as peptide net charge variation[1,2] or change in hydrophobicity/aromaticity[3] and varied sequence pattern[4] have been shown to affect the self-assembly properties of peptide sequences.

Many studies have pointed out the significance of solution pH and hence the net charge distributions on the self-assembly of varied peptide sequences. Caplan *et al.* (2000) showed that the oligopeptide KFE12 self-assembled into extended β-sheets only upon neutralization of excess charge. The author reported that the self-assembly of the peptide occurred when the intermolecular electrostatic repulsive forces were below van der Waals attractive forces[5]. In a subsequent work, the author showed that rational alteration of charged residues and the number of repeat units in the oligopeptide can be used to trigger the peptide to self-assemble under physiological salt concentration and pH[6]. In another study, Lopez *et al.* (2002) reported that the hexapeptides: KTVIIE, STVIIE and KTVIIT self-assembled into fibrillar structures only when the net charge on the peptide is ± 1[7]. Aromatic residues have been known to direct amyloid fibril formation in many amyloid inspired peptide sequences through π-stacking and aromatic hydrophobic interactions[8].

Different studies published by Gazit & co-workers have highlighted the central role of aromatic interactions in the self-assembly of short peptide sequences designed from different amyloidogenic proteins[9,10]. An important example is a hexapeptide -NFGSVQFV- and the shorter peptide –NFGSV-, derived from human hormone medin that self-assembled into amyloid-like fibrils. Alteration of the phenylalanine residue in the peptide sequences impaired the ability of the peptide to self-assemble[11]. Other examples include the diphenylalanine repeat (FF) peptide and phenylalanine oligopeptides inspired from $A\beta_{1-42}$ associated with Alzheimer's disease self-assembled into a range of structures including nanofibrils, nanowires, and nanotubes[12,13,14].

Interestingly, there are also self-assembling peptide sequences that are tolerant to subtle changes. Examples include the cyclic octapeptide Lanreotide, a synthetic analog of natural somatostatin hormone, which is tolerant to modification of lysine to D-lysine and mutation of aromatic residues[15]. In another work, Armstrong *et al.* (2011) showed that variation of phenylalanine residues at position 19 and 20 in the amyloid fibril forming $A\beta_{1-42}$ that is thought be important for self-assembly with leucine or isoleucine rather enhanced amyloid fibril formation[16]. Despite numerous examples that show that small variation in peptide sequence or change in the chemistry of the peptide effect peptide self-assembly, it is difficult to predict the impact of the change or mutation. Therefore, it is important to consider each peptide sequence and the effect of the change on a case by case basis.

4.2 Design strategy for alteration of Prx3 protein-β interface peptide

From previous studies, it was understood that short peptide sequences designed from protein β-interfaces self-assembled into extended amyloid fibril-like structures[17]. The assembly of such peptide sequences was stable over the pH range 5.0 to 8.0 (Chapter 3, Section 3.7).

In the specific case of Prx3 peptide, designed from the homodimer interface within the *bovine* Prx3 dodecamer[18], the histidine residue at the interface was found to be in close contact with serine and asparagine residue. Figure 4.1A shows the 3.3 Å resolution crystal structure of the *bovine* Prx3 dodecamer and the homodimer β-continuous interface. Measurement of bonding distance performed using pymol (www.pymol.edu) showed that the histidine residue at position 136 to be within hydrogen bonding distance to the neighboring serine residue (position 138) on the same β-strand and valine (position 139) and asparagine residues (position 140) on the complementary β-strand (Table 4.1 and Figure 4.1A).

Sequence alignment focused at the homodimer interface of 2-Cys peroxiredoxin 3 from eleven different species revealed that the histidine residue at position 136 in the crystal structure of *bovine* Prx3 (PDB 1ZYE) at the interface sequence is highly conserved. Aside from the histidine, the three hydrogen bonding partners: valine, asparagine, and serine residues were also conserved in all the species except for *S. cerevisiae*. Figure 4.2 B shows the sequence alignment generated using the Uniprot database and the ESPript sequence alignment software (http://espript.ibcp.fr).

Table 4.1: *Location and bonding distance between histidine and hydrogen bonding partners*

Location	Distance between residues
β9 strand	H136 to S138 (3.1 Å)
β9- β9’ strand	N140 to H136 (2.5 Å)
β9- β9’ strand	V139 to H136 (3.6 Å)
β9’ strand	H136 to S138 (3.0 Å)
β9’- β9 strand	N140 to H136 (2.6 Å)
β9’- β9 strand	V139 to H136 (3.5 Å)

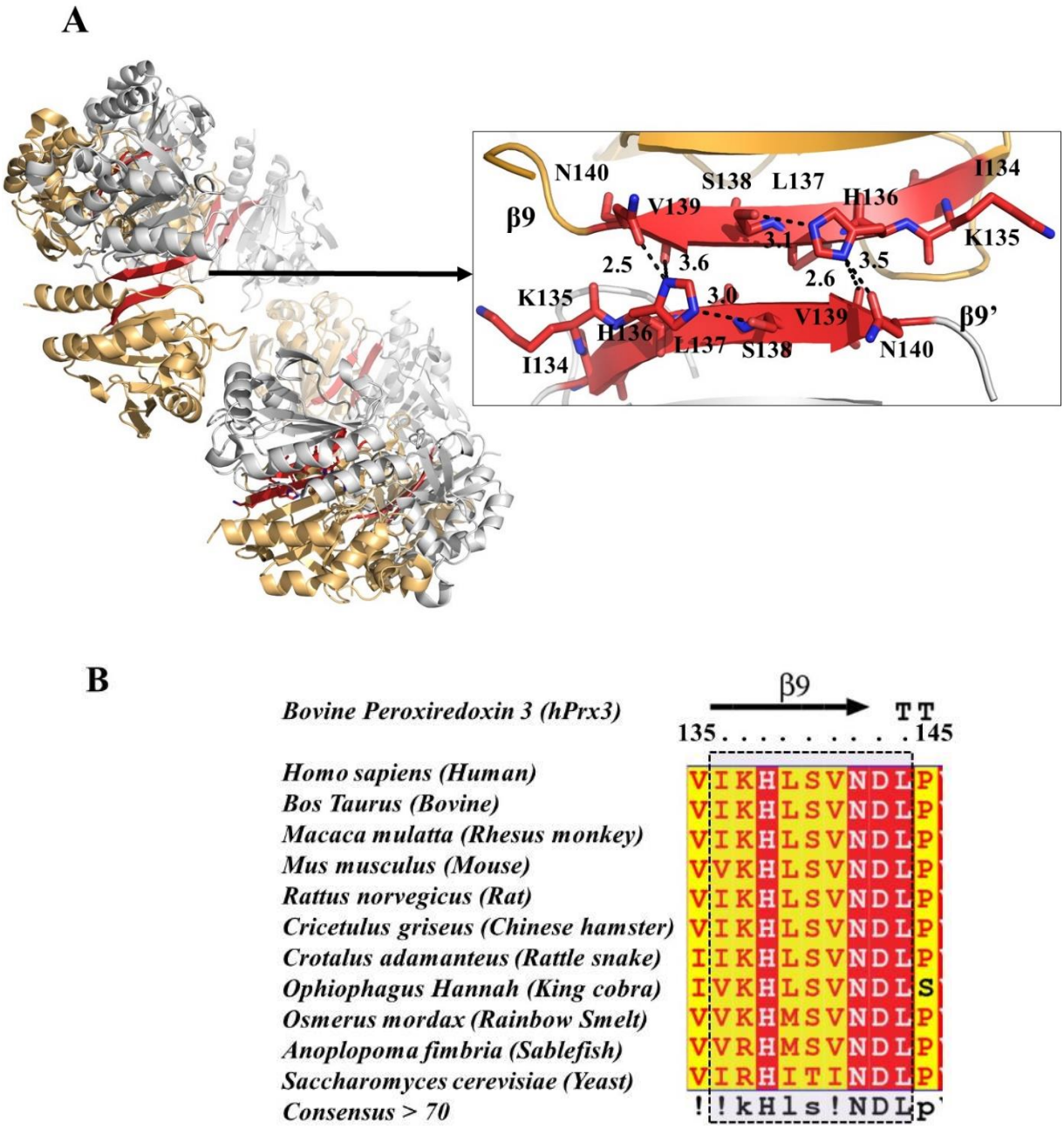


Figure 4.1: *Crystal structure and sequence alignment of peroxiredoxin 3: (A) cartoon model of the ring-shaped dodecamer toriod of bovine Prx3 (PDB:1ZYE). The inset shows the placement of the asparagine (N), serine (S) and histidine (H) residues within the homodimer protein-β interface and the hydrogen bonding distances are highlighted by black dotted line. (B) Sequence alignment of Prx3 from eleven different species (position:135-145) produced using Uniprot database and ESpript software, showing the region of high consensus. The black dotted box highlights the β-continuous interface sequence.*

The highly conserved positions of the histidine, serine and asparagine residues in the sequence making up the protein-protein homodimer interface suggested that they were necessary interactions at the interface. Therefore, it was posited that these interactions could be necessary for the self-assembly of the isolated peptide sequence derived from the interface (Prx3 peptide).

It is noteworthy that protonation of histidine residue as a function of pH is reported to be involved in the self-assembly process of different peptide and protein assemblies[19,20,21,22]. Histidine protonation due to pH change is described in the context of biomolecular switching in viral fusion proteins that facilitate the attachment of the virus particle to the host cell membrane[23]. More specifically, the interaction of the proximal histidine residue with hydrogen bonding partners such as aspartate and asparagine have been previously described in peroxidases and is concerned with important structural differences between non-animal (plant type) and animal heme peroxidases and dehydrogenases[24,25,26]. Carpena *et al.* (2009) showed that the imidazole side chain of histidine can exist in the negatively charged imidazolate form which is involved in hydrogen bonding with nearby asparagine and arginine residues in the high resolution (2.8 Å) crystal structure of human myeloperoxidase glycoform (*h*MPOm)[24].

In the case of peptides, hydrogen bonding between neighboring histidine and serine residues have been noted for peptide sequences that self-assembled into extended β-sheet based structures. In a recently published work, Valery *et al.* (2015) showed that the decapeptide triptorelin under alkaline conditions pH > 7.8 adopted a globular conformation in which the histidine residue in the peptide is involved in hydrogen bonding to the nearby serine residue and also involved in histidine-aromatic interactions. However, under acidic pH, such interactions are not present, and the peptide adopted an extended conformation and self-assembled into bundles of small twisted nanotubes[27].

These studies suggested that protonation and deprotonation of histidine residue is important for maintaining the higher order quaternary structure of peptide and protein assemblies. Thus, it was hypothesized that change of histidine residue in the Prx3 peptide sequence to a hydrophobic residue would limit the probable side chain hydrogen bonding interactions and therefore affect the ability of the isolated peptide sequence to self-assemble.

4.3 Variant Prx3 peptide sequences investigated

To study the effect of a specific sequence change on the self-assembly of the Prx3 peptide, hydrophobic variations to the histidine residue were investigated. The histidine residue, at position 3 in the Prx3 peptide sequence was replaced with the hydrophobic alanine (A) residue. Variation of histidine residue to aromatic tyrosine (Y), and phenylalanine (F) tryptophan (W) was performed to study the replacement of imidazole ring side chain with aromatic ring side chain. The final set of peptide sequences were seven residues long that corresponded to the β-continuous homodimer interface within the bovine Prx3 dodecamer[18] with variation of histidine residue at position 3.

The peptide sequences were capped at both ends: N-terminus acetylated and C-terminus amidated, to keep the overall charge on the peptide constant and comparability to the wild type peptide. The synthesized variant peptide sequences were obtained from Mimotopes Ltd. (Sydney, Australia) as a white lyophilized powder with purity > 95% confirmed by mass spectrometry and HPLC as provided by the manufacturer. Table 4.2 shows the parent peptide and variant peptide sequences explored in this work and their properties as calculated using the peptide property calculator {http://pepcalc.com/ (Innovogen, Sweden)}. The variant peptide sequences were named based on the variant residue introduced in place of histidine.

Table 4.2: *Prx3 peptide and designed variant peptide sequences*

Peptide	Sequence	Theoretical pI	Hydrophobic/ hydrophilic	Net charge at pH 7	Number of aromatic residues
Prx3 (WT)	Ac-IKHLSVN-NH_2	9.13	4/3	2.0	0
H3A	Ac-IKALSVN-NH_2	9.02	4/3	1.0	0
H3Y	Ac-IKYLSVN-NH_2	8.65	4/3	1.0	1
H3W	Ac-IKWLSVN-NH_2	9.02	4/3	1.0	1
H3F	Ac-IKFLSVN-NH_2	9.02	4/3	1.0	1

4.4 Characterization of variant peptide self-assembly in solution

From previous results, as described in chapter three (Section 3.4), it was understood that the parent peptide is soluble in organic solvents and, in water at low concentrations. Also, the peptide sequence self-assembled into extended nanostructures under hydrophobic conditions. Therefore, the variant peptide sequences were examined for solubility in organic solvents (HFIP, DMSO, and methanol), methanol in water mixes and water. Peptide acetate samples prepared as described in chapter two (Section 2.5.8), were found to be soluble in 100% HFIP, DMSO, and methanol in the 10 to 100 mg/mL range, similar to the parent peptide sequence.

For peptide acetates dissolved in methanol-water mix at 50%, 25%, 10%, and 5% methanol concentration the variant peptide acetates H3A and H3F were found to be less soluble in 25% methanol solution at concentration ≥ 5 mg/mL. Peptide acetates H3Y and H3W were, however, soluble with no visible change in solution appearance in 5% methanol solution up to 20 mg/mL. All the peptide acetates exhibited poor water solubility at dilute concentrations, similar to the parent Prx3 peptide, which suggested that the self-assembly of the peptides is favored under hydrophilic conditions.

Peptide acetate samples prepared at higher concentrations of 50 mg/mL for H3A and H3W and at 100 mg/mL for H3Y and H3F peptide variants formed turbid hydrogels over time at room temperature (Figure 4.2), consistent with the parent peptide.

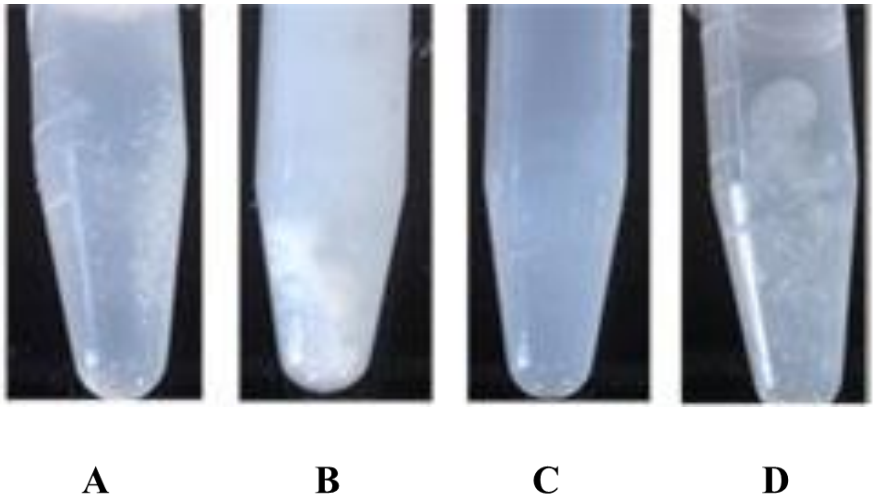


Figure 4.2: *Peptide acetate in water mixes of variant peptides: (A) H3A and (B) H2F at 50 mg/mL and (C) H3W and (D) H3Y at 100 mg/mL equilibrated for 24 hours at 4°C. The picture shows viscous hydrogels formed by self-assembly of variant peptide acetates that suggest an extensive network of fibrillar assemblies.*

The results suggested that variant peptide sequences behave similarly to the parent peptide sequence in solution. It is important to note that solution environment and concentration of interacting molecules in solution can affect the subtle equilibrium between the non-assembled monomer state and the self-assembled aggregate state. This is explained by the Le Chatelier's principle of general reaction of association, which states that the association and dissociation of molecules is a response triggered by the changes in the external solution environment [28,29,30]. Pressure effects arising from the interaction between water molecules and the exposed hydrophilic groups of amino acid residues affect this equilibrium and drive self-assembly[31,32].

Turbid hydrogels analyzed by cross-polarized light microscopy exhibited birefringence, with liquid crystalline textures observed for all the four variant peptide acetates in water in the 50 to 150 mg/mL range (Figure 4.3).

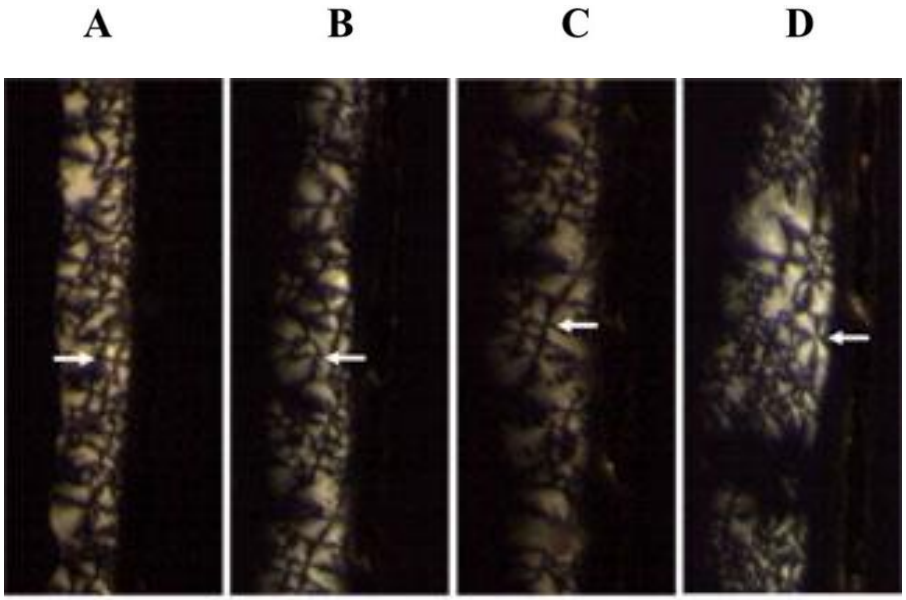


Figure 4.3: *Birefringence of variant peptide hydrogels observed under polarized light with the cross-polarizers at (± 45°). Optical textures of square columnar phase were observed for the peptide acetates in water at 50 mg/mL for H3A (A) and H3F (B), 100 mg/mL for H3W and 150 mg/mL for H3Y, conditioned between two glass slides. The arrows point to one of the highest points of curvature of the specific textures.*

Birefringence under cross-polarizers is reflective of extended well-ordered assemblies consistent with that observed for amyloid-like fibrils[33]. The presence of liquid crystalline textures, consistent with the parent peptide (Chapter 3, Section 3.4) suggested that the variant peptide sequences self-assembled into large fibrillar structures in solution[34,35]. The variant peptide sequences exhibited a propensity for β-sheet secondary structure conformation consistent with the parent peptide, shown by the minimum at ~215 to 220 nm assigned to the β-sheet secondary structure[36], for dilute peptide acetates in water analyzed by CD spectroscopy (Figure 4.4).

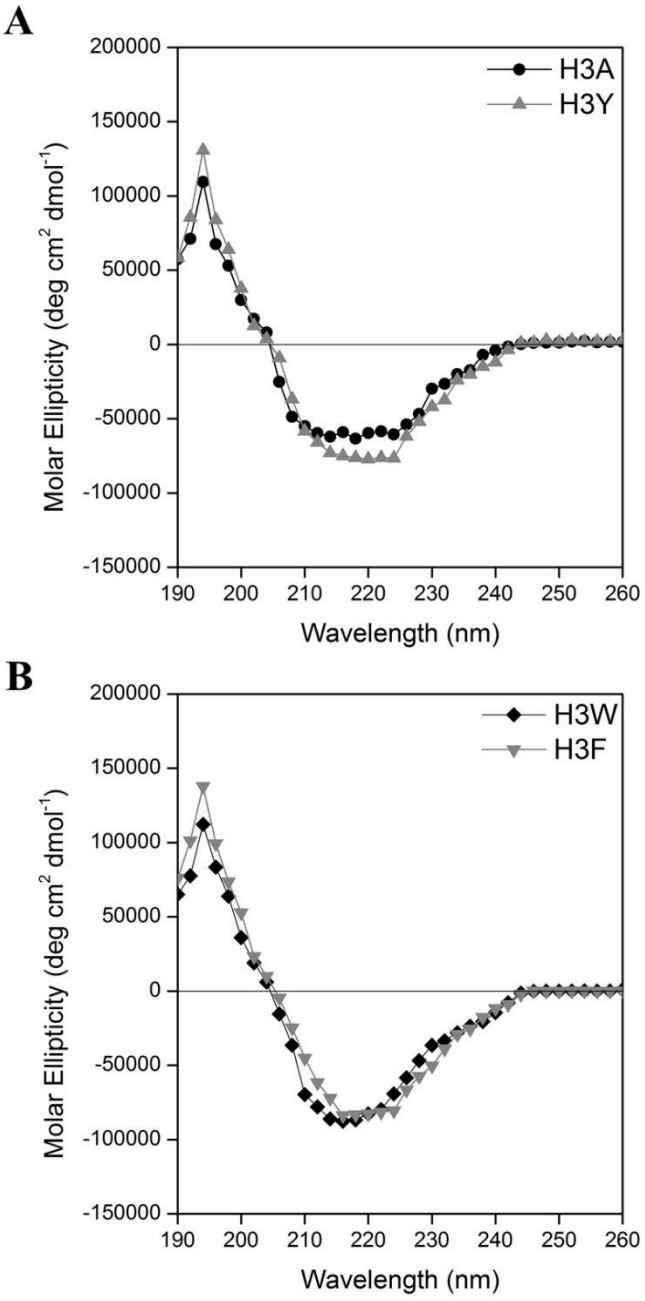


Figure 4.4: *Circular dichroism spectra of variant peptide acetates: (A) H3A and H3Y and (B) H3W and H3F in water at 0.01 mg/mL in the far U.V. range (190 - 260 nm).The peptide acetate in water samples was equilibrated for 24 hours 4°C and degassed before analysis.*

The result suggested that introduction of hydrophobic variation of histidine in the wild type Prx3 peptide sequence does not substantially alter the propensity of the peptide to exist in β-sheet secondary structure conformation.

4.5 ThT binding assays

ThT binding assays on the parent Prx3 peptide showed that the peptide self-assembled in solution in a reversible manner, as described in chapter 3. Therefore, self-assembly of the variant peptides in solution was monitored using ThT fluorescence as a function of concentration and time duration. Peptide acetate samples for each of the four peptide variants and control parent Prx3 peptide prepared in water (w/w) were recorded at 24 hours after equilibration as described in chapter 2 (Section 2.5.6).

All the four variant peptide sequences exhibited strong, though, to a lower extent compared to the parent peptide but consistent with the presence in each of the variant peptides of the ordered intermolecular β-sheet structures. ThT binding curves generated by plotting the derivative of ThT binding (dt/dc) versus peptide acetate concentration (mg mL^{-1}), exhibited a linear derivative curve for the alanine variant H3A and the phenylalanine variant H3F, similar to the parent peptide. In the case of the aromatic variants H3W and H3Y, the peptide variants exhibited a sigmoid growth type derivative curve.

These results suggested that the H3A and H3F variants exhibit a single state model of self-assembly in which the self-assembled peptide β-sheets dominated in solution. The replacement of the histidine residue in the parent peptide sequence to the aromatic tryptophan or tyrosine residue, however, exhibit a two state self-assembly model. The model is consistent with the presence of a critical assembly concentration, a state of transition from prefibrillar oligomers to self-assembled β-sheet structures[37] (Figure 4.5). The results suggest that the introduction of a tyrosine residue with an extra hydroxyl group compared to phenylalanine and the tryptophan residue with the large aromatic side chain results in the change of self-assembly equilibrium.

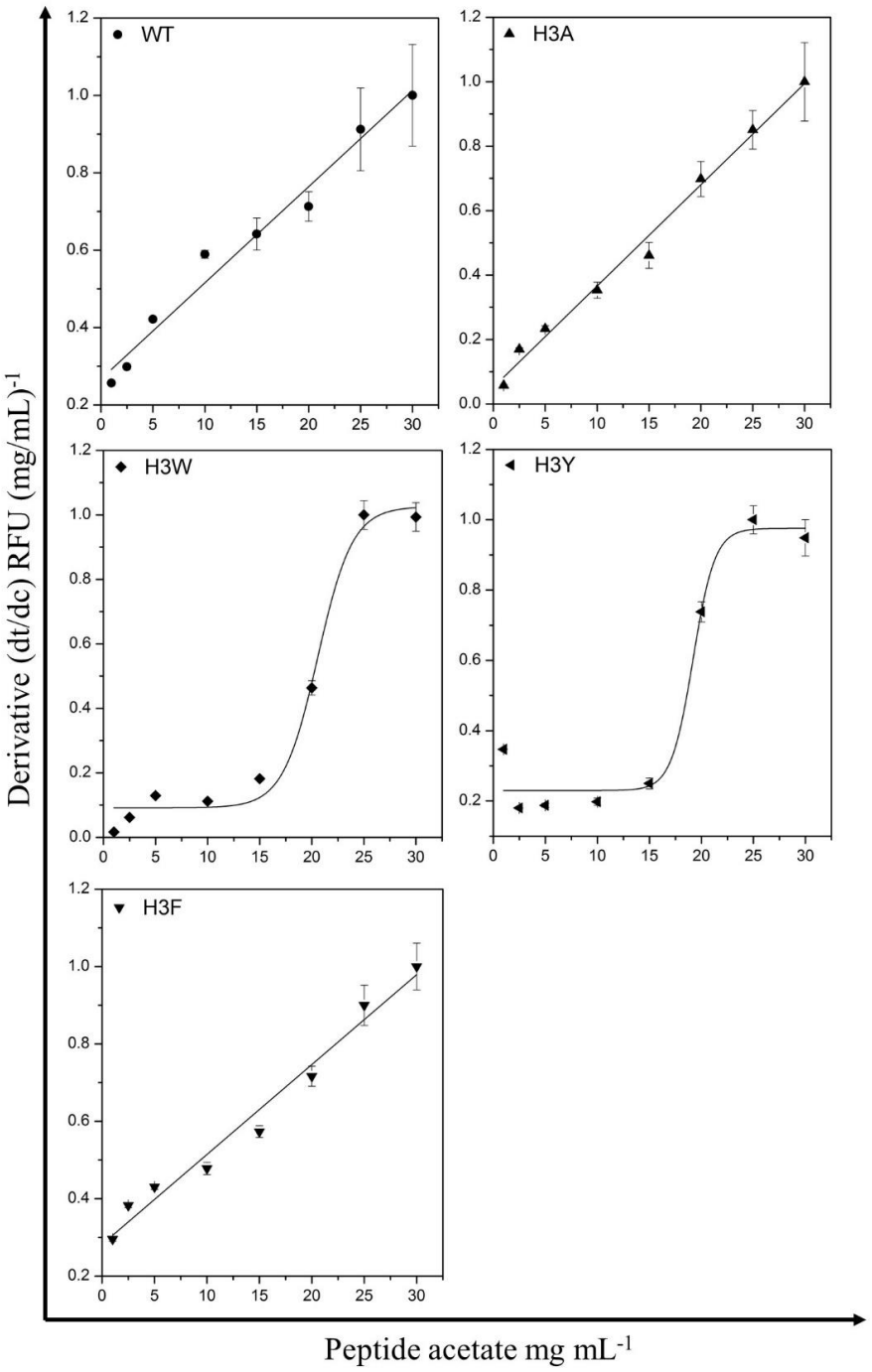


Figure 4.5: *ThT binding assays by dilution for Prx3 peptide (WT) and peptide variants. Fluorescence intensity values recorded 24 hours after dilution for samples equilibrated for 24 hours. The derivative of Prx3 peptide, H3A variant and H3F peptide variant fitted to a linear function of y = Ax + B. The derivative of H3W peptide variant and H3Y peptide variant fitted to a sigmoid growth function of* $y=B + (A + B)/(1 + e^{(x-x0)/dx})$*.The error bars correspond to the standard deviation for triplicate sample preparations.*

This behavior of the variant peptide sequences can be explained by the Le Chatelier principle which states that a small change in the environment, such as temperature, pressure or solute molecules observes a resulting change in the subtle equilibrium between the monomer state and the self-assembled or aggregated state[38,39]. The solute effect on the equilibrium is due to the preferential solvation of each group exposed on the surface with respect to solvent and co-solvent molecules[29,30].

4.6 Kinetics of peptide self-assembly in solution

The kinetics of self-assembly of the variant peptide sequences in solution was examined by ThT binding: (i) at three different concentrations and (ii) at different time points. Peptide in acetate mixes prepared at 1 mg/mL, 5 mg/mL and 25 mg/mL were monitored for ThT binding immediately after dissolving in water and at subsequent time intervals over 48 hours for peptide solution samples stored at 4°C. The ThT fluorescence intensity fitted to a single exponential function is shown in figure 4.6. ThT binding recorded immediately after dissolving the peptides in water showed enhanced ThT fluorescence intensity for all four variant peptide acetates in water at 25 mg/mL, although to a lower extent compared to the parent Prx3 peptide.

ThT binding recorded over a 48-hour time course for the peptide acetate solutions showed an increase in the ThT binding intensity over time for all four peptide variants consistent with the parent peptide. This suggested that the development of extended peptide β-sheet structures occurs over time and follows a kinetic profile that is effectively a single exponential[40,41]. The rapid self-assembly of the peptide acetates in water at higher concentration (25 mg/mL) into extended β-sheets that bind to ThT dye can be explained by the presence of β-sheet nuclei that affect the dynamic equilibrium between the peptide monomers and the self-assembled peptide β-sheet[42].

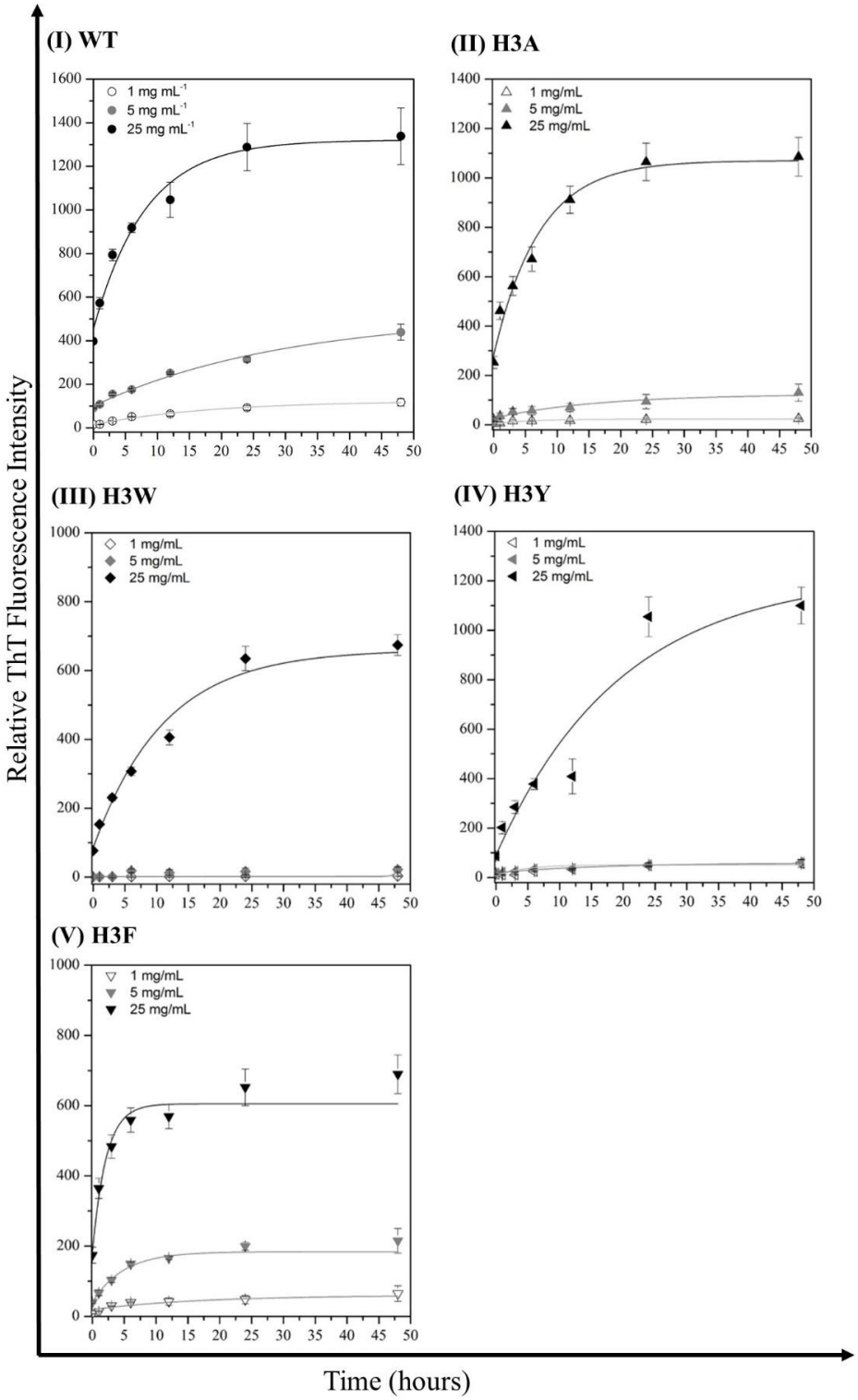


Figure 4.6: *ThT binding assays for peptide acetates in water recorded at different time-points for samples equilibrated at 4°C at three different concentrations. The curves fitted to a single exponential function of* $y = A + B*e^{(-kx)}$. *The error bars correspond to the standard deviation for triplicate sample preparations.*

The substantial difference observed in ThT binding between peptide acetates at 25 mg/mL compared to peptide acetate solutions with a five-fold lower concentration of the peptide at 5 mg/mL and 1 mg/mL respectively, suggested a nucleation-dependent assembly. This mode of self-assembly has been previously described for peptides and proteins that form amyloid fibrils[43,44].

The absence of any substantial ThT binding for dilute peptide acetate in water samples for the aromatic variants H3W and H3Y suggested a higher critical assembly concentration for the variant peptides. The results, obtained for the H3Y peptide variant are quite distinctive since H3F and H3Y are nearly identical with the only difference being the phenolic hydroxyl in the latter peptide. The H3Y peptide is less hydrophobic than the H3F peptide variant; this does seem to contribute to the difference in the rate of aggregation observed between the two peptide sequences. The difference in the aggregation of the lesser hydrophobic H3Y variant peptide is explained by the possibility of the hydroxyl group that is additional in the tyrosine amino acid to be involved in hydrogen bonding interactions with other residues. Also, this limits the specific directionality necessary for aromatic interactions, which is supported by negligible ThT binding observed for the H3Y peptide variant at dilute concentration in water.

A similar observation is apparent for the specific variation of the histidine residue in the parent peptide to the aromatic tryptophan residue (H3W). This behavior of the variant peptide sequence can be attributed to the large aromatic side chain of the tryptophan residue that causes in steric hindrance, which is, however, overcome at higher concentration of the peptide acetate in water.

Apparent rate constants for aggregation (k_{agg}) were calculated from the best fit of the single exponential[45,46], shown in figure 4.7. The k_{agg} values calculated for the peptide variants at the highest concentration (25 mg/mL) were similar to the parent Prx3 peptide (0.20 ± 0.04), except for H3F peptide variant which has a twofold higher k_{agg} values (0.47 ± 0.1). The larger k_{agg} observed for the H3F variant can be explained by the increase in hydrophobicity of the peptide sequence which results in the burying of the phenylalanine side-chain away from the water molecules and the probable π-π interactions that stabilize the β-sheet[10,12].

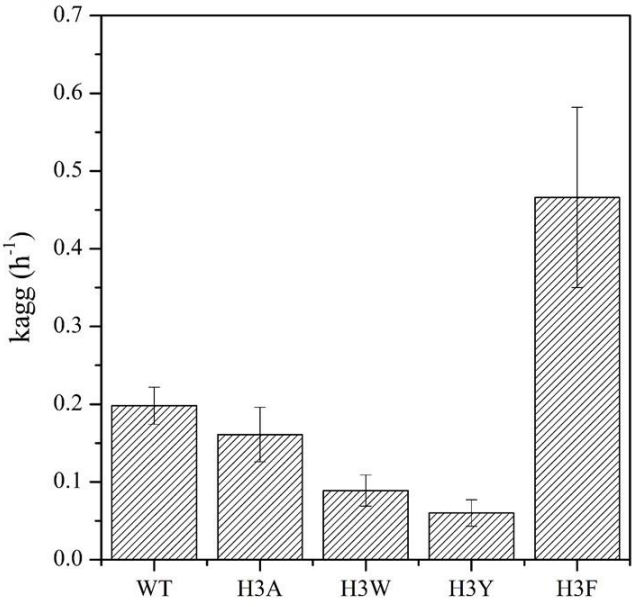


Figure 4.7: *Apparent rate constants of aggregation (k_{agg}) for the peptide acetates in water at 25 mg/mL, obtained from the best fit to the data shown in figure 4.6. Errors correspond to the standard errors of the means of three independent experiments.*

Subtle change in peptide chemistry though sequence modification can have varied effect in the self-assembly process. The results presented here, however, show that introduction of hydrophobic variation in the Prx3 peptide sequence through replacement of histidine residue does not substantially affect the self-assembly of the peptide into extended β-sheet based structures.

Hydrophobicity of the peptide sequence is known to facilitate aggregation of the peptide into extended β-sheet structures[47,48,49]. Also, aromatic stacking of the peptide sequence through π-π interactions, especially contributed by the phenylalanine is known to drive fibrillation of various peptide sequences[50,51,52,53]. The specific hydrophobic interactions and the possible aromatic stacking between the aromatic amino acid residues, therefore explains the self-assembly of the variant peptide sequences consistent with the parent Prx3 peptide. The small difference observed in the aggregation kinetics for the variant peptide sequences can be attributed to the shift in the aggregation equilibrium due to sequence modification. This change in the dynamic equilibrium of self-assembly of the hydrophobic peptide variants is explained by the pressure effects resulting from preferential solvation and the possible burying of the hydrophobic residues with respect to water molecules[31]. Overall the results suggested that the self-assembly of peptide variants followed kinetics profile characteristic of cooperative processes[54,55].

4.7 TEM characterization of variant peptides

To study the morphology of the β-sheet structures formed by the variant peptide sequences, negatively stained (uranyl acetate) grids of peptide acetates in water in the range 10-50 mg/mL were analyzed by TEM (Section 2.1.7). All four peptide variants self-assembled in solution to form nanostructures with extended morphology with bundles of nanotapes observed for H3A, individual nanofibrils and wide nanotapes for H3Y, individual nanofibrils for H3W, and H3F peptide variant consistent with the parent Prx3 peptide (Figure 4.8). The heterogeneous mix of peptide fibrils and wide peptide nanotapes observed for the variant peptide H3Y was suggestive of polymorphism. Polymorphism is known characteristic of amyloid fibrils[41,56] and has also been described in the context of amyloid-like fibril assemblies formed by short peptide sequences[57,58,59].

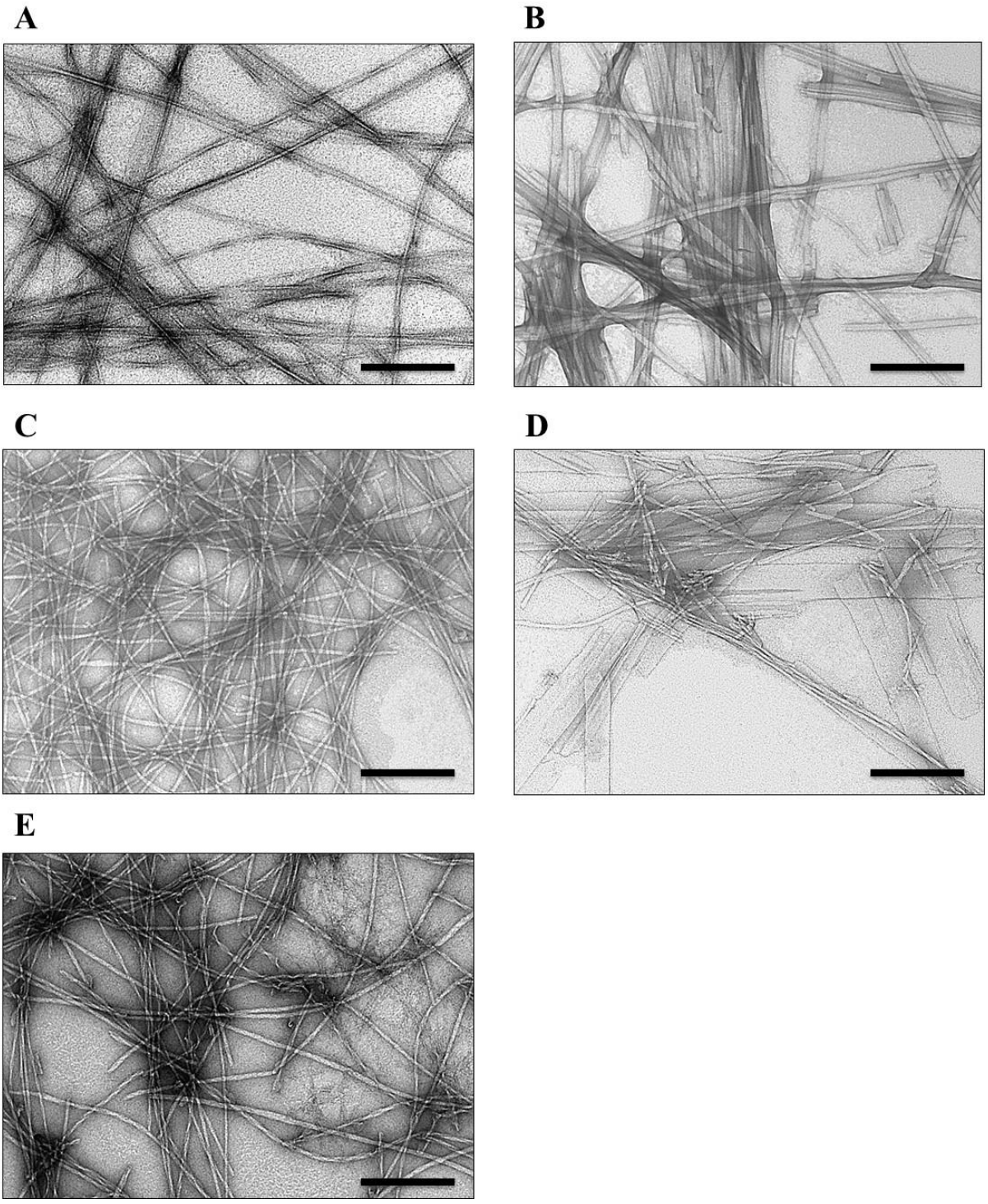


Figure 4.8: *Representative TEM micrographs variant peptide acetates: Prx3 peptide (A), H3A peptide (B), H3W (C), H3Y (D) and H3F (E) in water at 50 mg/mL equilibrated for 24 hours. (Scale bar is 200 nm)*

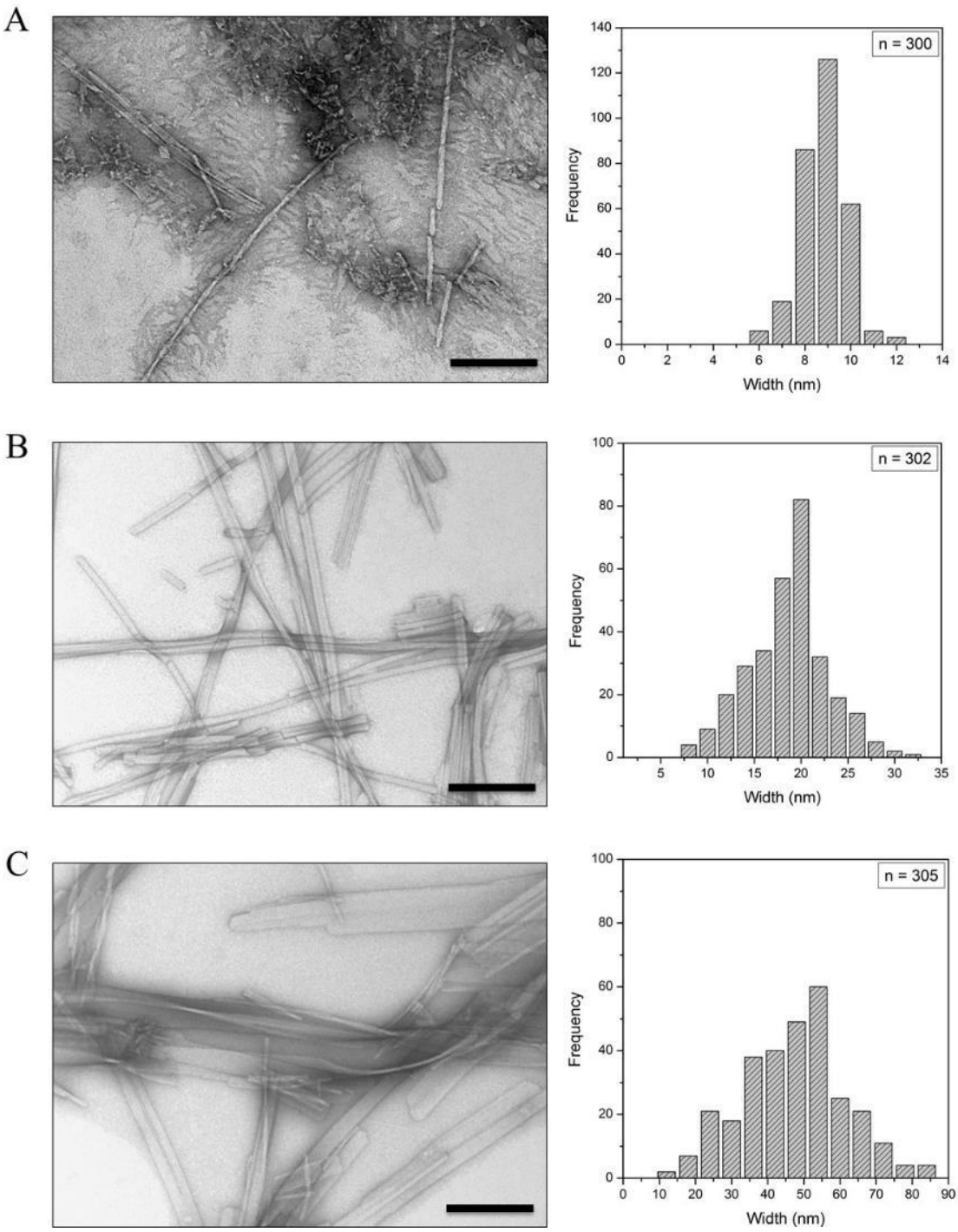


Figure 4.9: *Representative TEM micrographs of H3Y peptide acetate in water at 15 mg/mL (A), 25 mg/mL (B), 50 mg/mL (C). The micrographs were taken after equilibration of the sample for 24 hours. The histograms shown the size distribution of the widths of the fibers observed in each case, collected from ≥ 300 particle measurements done using ImageJ. (Scale bar is 200 nm).*

Measurement of polymorphic nanotapes observed for H3Y variant peptide revealed peptide tapes with an average width of 50 nm at 50 mg/mL. Dilute peptide acetate solutions at 25 mg/mL revealed narrow tapes with an average width of 20 nm and thin peptide fibrils observed at 10 mg/mL were on an average 10 nm wide (Figure 4.9). This type of size distribution of peptide nanotapes suggested the ability of the variant peptide sequence to self-assemble into extended fibrillar nanostructures of variable morphologies.

The hallmark of many different peptide sequences that self-assemble into amyloid-like fibrils is the ability to self-assemble into fibrils of different morphologies. It is suggested that the structural polymorphism observed for the diverse self-assembling peptide sequences is due to the different packing of the peptide chains within the extended fibrillar structure[56,57,60,61]. Fibril polymorphism is reported to be controlled by growth conditions and also seed fibrils with a particular morphology that induces the formation of fibrils with same morphology in the sample[62,63]. In a recent atomic level resolution structure of transthyretin peptide fibrils, sideways stacking of β-sheets facilitated through terminal side-chain interactions resulted in a different number of β-sheets laterally arranged in the extended fibril architecture[58,64].

4.8 Small angle X-ray scattering of variant peptides

To study the repetitive molecular organization within the extended nanostructures observed for the parent peptide and the variant peptide sequences by TEM, samples of peptide acetates were analysed for SAXS patterns. Concentrated samples of the variant peptide acetates in water in the 50-200 mg/mL range and the parent peptide at 50 mg/mL in water and pH 8.0 imidazole buffer were analyzed at the SAXS beamline (section 2.6.9). The SAXS patterns of the variant peptide acetates in water showed the characteristic Bragg peak at q = 1.3 $Å^{-1}$, consistently with the parent Prx3 peptide acetate in water and imidazole pH 8.0 buffer (Figure 4.10).

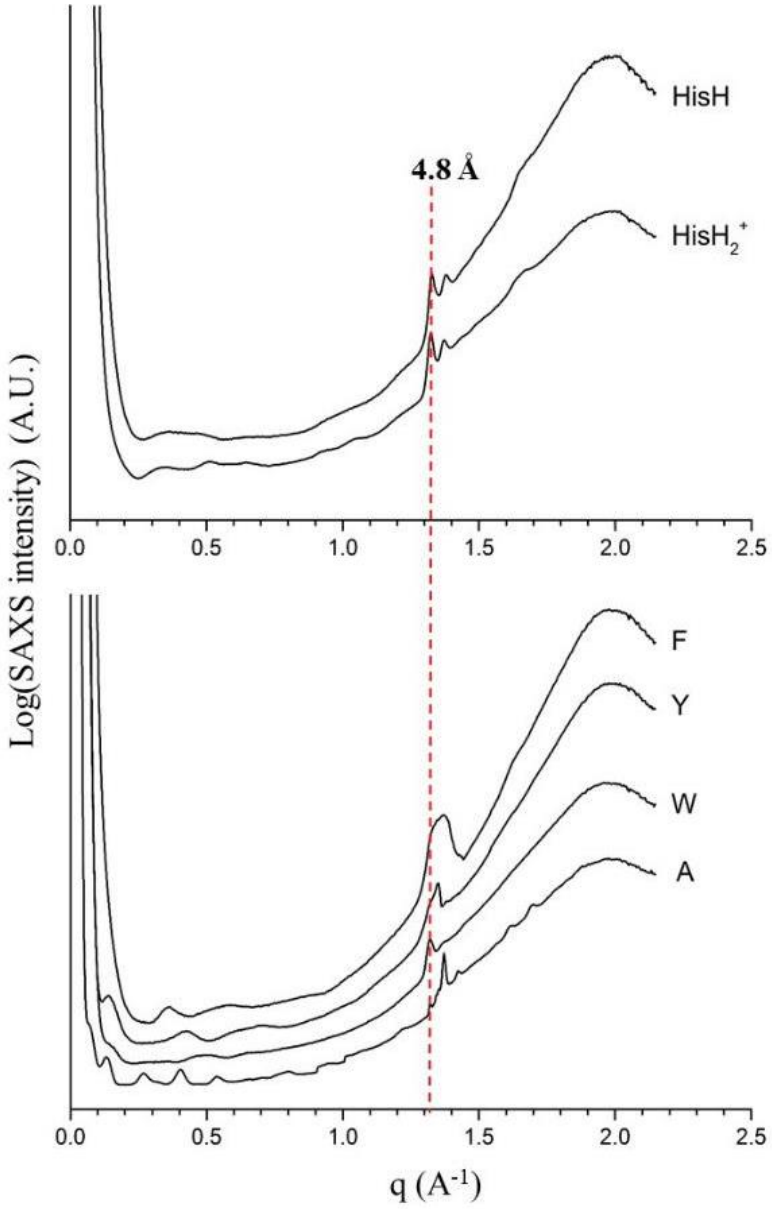


Figure 4.10: *SAXS patterns recorded at room temperature for peptide acetates in water & in pH 8.5 imidazole buffer for the Prx3 peptide (HisH and $HisH_2^+$) and variant peptide acetates in water (F, Y, W, and A). The vertical dashed red line points out the reflection at 1.3 $Å^{-1}$ visible for all the peptide sequences.*

The Bragg's peak observed for the wild type peptide and the variant peptide corresponds to the characteristic meridional reflection that arises from the spacing between the adjacent hydrogen bonded β-strands along the fibril axis. This type of SAXS pattern has been previously described for varied β-sheet type self-assembling peptides sequences and amyloid fibrils[65,66,67].

The lesser defined broad scattering observed at the small angles below 0.5 $Å^{-1}$ for both wild type peptide, and variant peptides was attributed to diffused scattering of the samples. The well-defined reflections, however observed for the H3A peptide at smaller angles can be attributed to the formation of micro/nano-crystals in concentrated peptide acetate sample that contribute to more defined X-ray scattering. Such patterns have been previously described for fibrillizing peptide sequences and correspond to an equatorial reflection thought to arise from the anisotropic form factor of nanofibrils perpendicular to the fibril axis [54,68,69]

Overall the SAXS patterns showed that the molecular organization of the variant peptide sequences is similar to the parent Prx3 peptide, consistent with the repetitive organization of the β-sheets over long distances within the of the extended peptide nanostructures.

4.9 Fourier transform infrared spectroscopy of variant peptides

The arrangement of the peptide strands within the extended β-sheet assemblies of the peptide sequences was examined by FTIR microscopy at the Australian Synchrotron Facility (Section 2.6.10).

The FTIR spectrum recorded for concentrated peptide acetate in water (10% w/w) focused on the amide I vibrations (1600 – 1700 cm^{-1}) which correspond to the vibrational stretching of the backbone carbonyl bonds[70]. The IR spectra obtained for the variant peptides and the parent peptide exhibited three well-defined infrared vibrations at ~1620 cm^{-1} , 1640 cm^{-1} , between 1673 to 1685 cm^{-1} and lesser defined vibration around 1660 cm^{-1} (Figure 4.11).

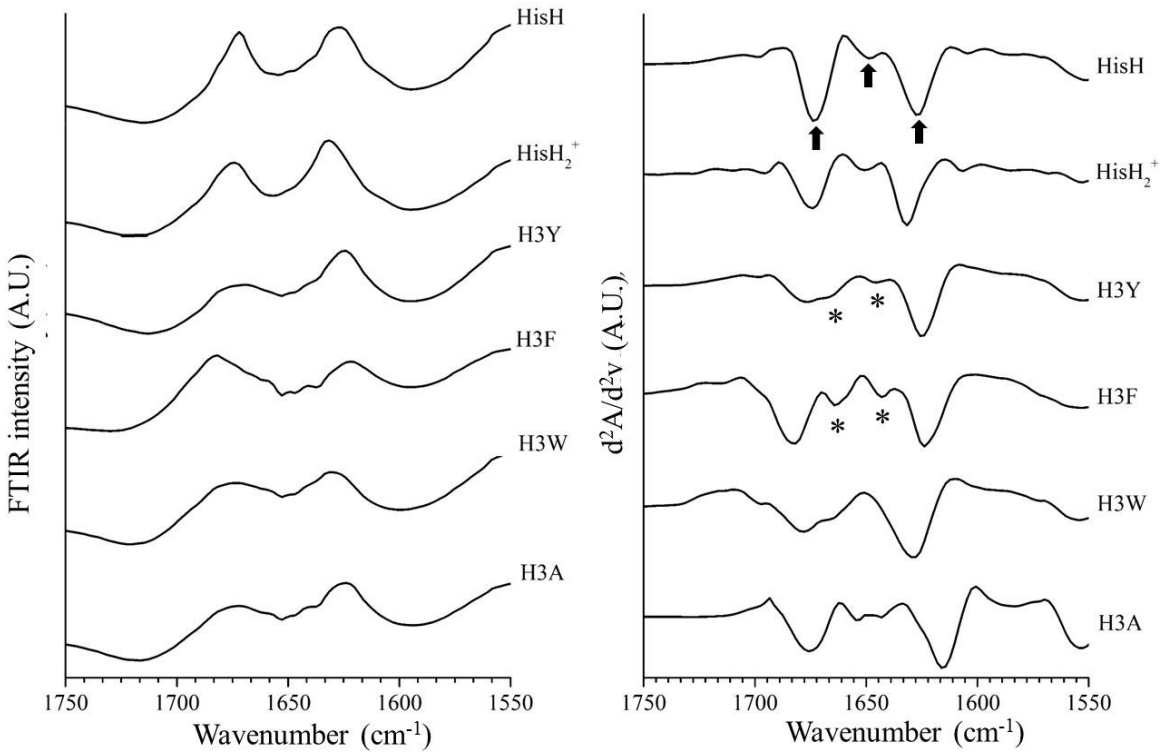


Figure 4.11: *IR absorbance spectra for Prx3 peptide acetate in water & pH 8.5 imidazole buffer (HisH and $HisH_2^+$) and variant peptide acetates in water (F, Y, W and A) recorded at room temperature. The black arrows show the distinct absorption peaks around 1620 cm^{-1}, 1640 cm^{-1} , 1650 cm^{-1} and 1680 cm^{-1} and the "*" points to the lesser defined peaks at 1640 cm^{-1} and 1660 cm^{-1}.*

The presence of the combination vibrations at ~1620 cm^{-1} and at 1673-1685 cm^{-1} was assigned to the antiparallel β-sheet secondary structure in accordance with the literature [70,71,72]. The vibration around 1640 cm^{-1} was assigned to random conformations[70,73,74,75]. The fourth vibration around 1660 cm^{-1} was more pronounced for the aromatic peptide variants and assigned to a turn in thesecondary structure[76,77]. This combination was in agreement with the literature. The pair of vibrations observed around 1620cm^{-1} and 1680 cm-1 in each of the cases was consistent with the presence of antiparallel β-sheets, in the core cross-β architecture reported for the amyloid fibril fold[78,79,80]. Table 4.3 details the different vibrations observed for the parent Prx3 peptide and the peptide variants.

Table 4.3: *IRM amide I vibrations of the Prx3 peptide and variant peptides H3F, H3Y, H3A, and H3W. Wavenumbers taken from the 2nd derivatives of the IRM spectra and the corresponding secondary structures.*

Peptide	Antiparallel β-sheet (cm^{-1})	Random coil (cm^{-1})	Antiparallel β-sheet (cm^{-1})
HisH (WT)	1627	1650	1674
$HisH_2^+$ (WT)	1629	1651	1674
H3Y	1625	1646	1676
H3F	1623	1643	1682
H3W	1627	1654	1678
H3A	1620	1646	1675

The small variation observed in the spectra between the control peptide Prx3 and peptide variants could be due to the small variation in the repetitive symmetries in the extended β-sheet assemblies[77]. The amide I vibrations here described are consistent with literature [81,82,83] and suggestive of the anti-parallel conformation of the peptide in each of the four cases. Altogether the IRM results demonstrated that the peptide variants adopt a well-defined antiparallel β-sheet conformation in the extended nanostructure similar to the wild type Prx3 peptide.

4.10 Design of *h*Prx3 protein-β interface muteins

Human peroxiredoxin 3 (*h*Prx3) belongs to the family of thioredoxin-dependent peroxidase enzymes that fix peroxides in the cells[84] and exists as a stable dimer through the β-continuous interface[85]. This homodimer conformation is crucial for the organization of protein into the hierarchical toroid shaped structures including octamer[86], decamer[87], dodecamer[18]. The homodimer interface of the *h*Prx3 protein is symmetrical and contributed by two antiparallel β-strands (β9-β9') in complementary non-covalent close contacts (Figure 4.1A). An important feature of the β-interface motif is the amphipathic nature with a solvent accessible hydrophilic face and a hydrophobic face. The complementary interactions along the interface motif contribute to the stable dimerization of the protein units.

The histidine residue within the Prx3 peptide sequence derived from the homodimer interface of peroxiredoxin 3 protein was thought to be an important residue for self-assembly. The Prx3 peptide corresponding to the homodimer β-continuous interface within the Prx3 dodecamer has been shown to reversibly self-assemble into extended β-sheet based nanotapes[17]. The assembly of the peptide into extended β-sheet nanostructures was resilient to specific hydrophobic variations of the histidine residue. This indicated that the backbone hydrogen bonds were well preserved within the extended β-sheet self-assembly of the parent peptide and variant peptide sequences.

Short peptide strands have a larger degree of freedom to associate through backbone hydrogen bonding into extended β-sheets[88] than the parent interface, where this freedom is restricted in the 3-dimensional native fold of the protein. Therefore to study the effect of the same variation at the homodimer interface in the parent protein, single residue mutants of *h*Prx3 protein were investigated.

4.10.1 Expression and purification of *h*Prx3 muteins

The *h*Prx3 protein that bears 93% sequence homology with *bovine* Prx3 and 100% sequence homology at the homodimer interface was used for this work. Hydrophobic muteins of *h*Prx3 protein containing the specific mutation of the histidine residue at the homodimer interface were produced through site-directed mutagenesis. The protein muteins prepared as recombinant constructs comprised with the specific mutation of the histidine residue at the position 136 and the 6*His affinity peptide tag at the N-terminal end (Section 2.1.5).

The wild type *h*Prx3 and protein muteins were expressed as N-terminus 6*His-tagged protein constructs in chemically competent *E. coli* BL21 DE3 *Rosetta* cells cloned with a *pET-151-d-Topo* plasmid containing the gene for the *h*Prx3 protein[89]. Small scale flask cultures were trialed to ascertain the best conditions for a decent yield of protein, which were then used for large-scale expression of protein. A single colony of cells containing the specific *h*Prx3 WT and the mutein gene was used to inoculate a 10 mL starter culture grown overnight at 37°C. The starter culture was used to inoculate a 400 mL flask culture containing appropriate antibiotics grown at 37°C.

Protein expression was achieved through IPTG induction (400 µM) after $O.D._{600}$ was reached (Section 2.5.2.1). The expressed protein was subsequently extracted by sonication (Section 2.5.5.1) and purified by immobilized metal affinity chromatography (Section 2.5.5.4). Fractions with purified protein were judged by SDS-PAGE (Section 2.6.2), pooled and further subjected to preparative size exclusion chromatography (Section 2.5.5.5), as shown in Figure 4.12A and B for wild type *h*Prx3.

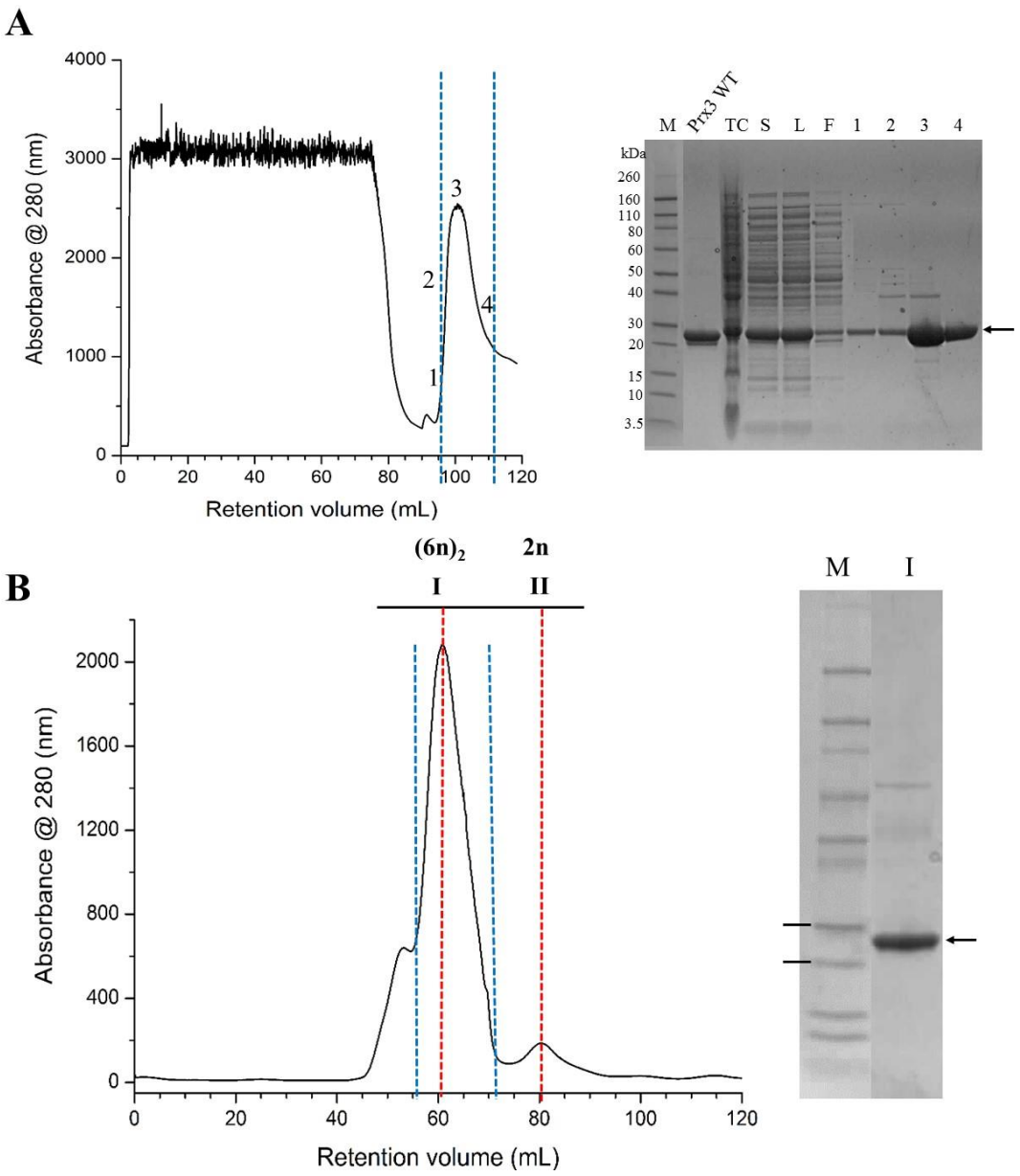


Figure 4.12: *Purification of hPrx3: (A) Immobilized metal ion chromatography (IMAC) profile and SDS-PAGE gel of hPrx3 (WT). (B) The sample was run on a GE HiTrap IMAC FF column at a flow rate of 1.0 mL/min equilibrated with Hepes, imidazole pH 8.0 buffer.*

*Note: Blue dashed lines represent the eluted fraction of the 6*His-tagged protein. Reducing SDS-PAGE gel shows the different steps of the purification process: TC; total cell lysate, S; soluble fraction, L; IMAC load, F; flow through, Eluted fractions (1,2,3 and 4). (B) Preparative SEC profile and reducing SDS-PAGE gel of eluted fraction II.*

The conditions of expression and purification of muteins were similar in each case to that of the wild-type *h*Prx3 protein to ensure uniformity and valid comparison between them (Figure 4.13).

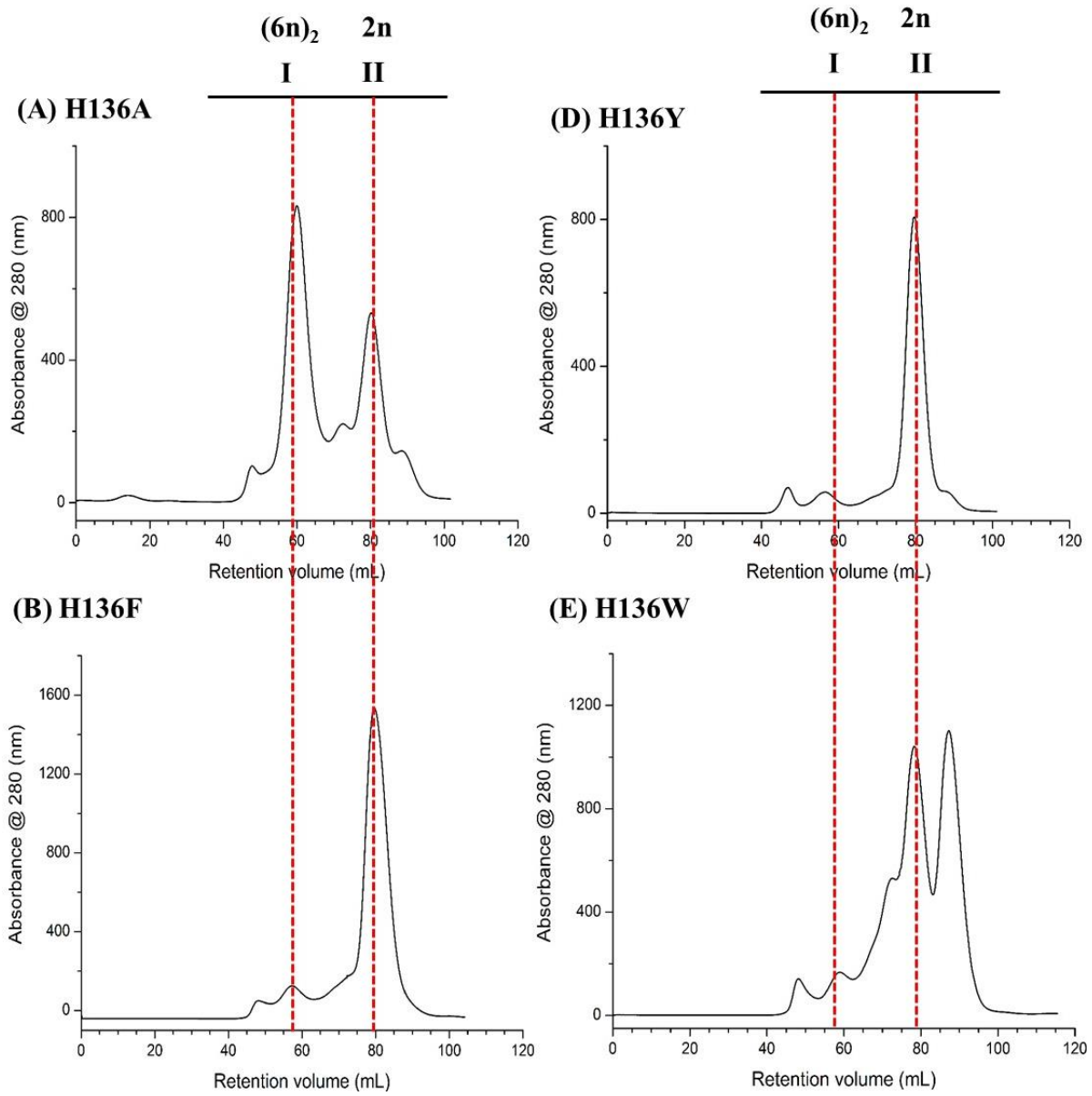


Figure 4.13: *Preparative SEC profiles of hPrx3 muteins. All the protein muteins were analyzed in HEPES pH 8.0 buffer on a Superdex 200 16/600 PG column at a flow rate of 1.0 mL/min. The Roman numbers indicate the four prominent protein species separated through chromatography. Dashed lines represent the dodecamer protein oligomer* $(6n)_2$ [90] *and protein dimer (2n).*

The wild type 6*His *h*Prx3 protein eluted mostly as a single peak at 60 mL retention volume, which corresponded to the toroid shaped single ring form of the protein. The lesser defined earlier eluting shoulder peak at ~ 50 mL and the later eluting peak at 80 mL corresponded to the high molecular aggregate and the dimer form of the protein respectively. This type of distribution by size exclusion chromatography is usually encountered for the *h*Prx3 protein in solution[91]. The muteins eluted as a mixer of oligomeric species, with the peak at 80 mL corresponding to the dimer form of the protein that is pronounced for H136F and H136Y muteins. For the specific change of the histidine residue to the aromatic tryptophan residue with a larger double ring side chain, the homodimer interface within the *h*Prx3 dodecamer appeared to have been weakened, consistent with the presence of the monomer form of the protein in solution.

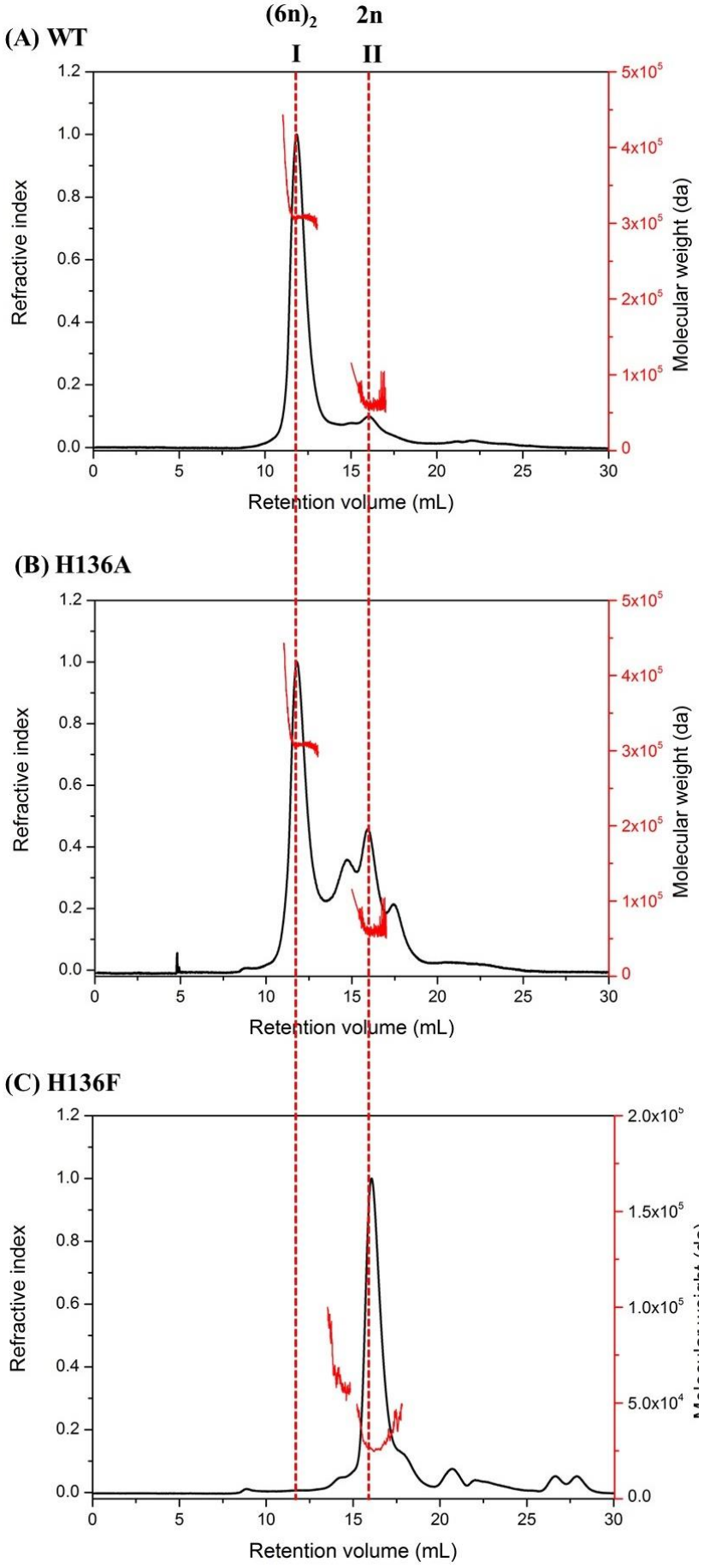
The wild type *h*Prx3 is a typical 2-Cys peroxiredoxin with a peroxidative thiol (C_P) residue and that an adjacent thiol cysteine (resolving cysteine C_R). The central β-sheet within the protein is oriented to give rise to a stable parallel β-sheet organization. The exposed β9 motif of one protein monomer associated with the complementary β9' motif of the other monomer protein to form a continuous β-sheet interface. The protein dimer locked through the B-interface, stabilized by the different non-covalent interactions across the β9- β9'motifs. The amino acid sequence of the interface motifs within protein oligomers are evolutionary fine-tuned to drive self-assembly of the protein monomer units in a regulated and controlled manner. The results here show that introduction of a hydrophobic mutation at the homodimer interface (B-interface) of *h*Prx3 protein by replacement of histidine residue do not have a substantial effect on the dimerization of the protein. For the specific mutation of the histidine to tryptophan, the monomer form of the protein is present in higher amounts relative to the protein dimer however, the data does not reveal whether this is at the expense of the dimer formation.

4.10.2 Solution state characterization of *h*Prx3 and dimer interface muteins

Analytical size exclusion chromatography coupled with a static light scattering detector (Section 2.6.3) allowed estimation of molecular weights of the different protein oligomers in solution. The size exclusion purified wild type *h*Prx3 protein and interface muteins at 1 mg/mL (load concentration) was subjected to in-line static light scattering. The molecular weight estimation was performed using a BSA standard protein injected before the sample protein and at the end of the run. Figure 4.14 shows the SEC-SLS profiles of the wild type *h*Prx3 and muteins.

The control wild-type *h*Prx3 protein showed mainly a single peak at retention time of ~12 minutes, which was estimated to be of ~ 300 kDa (I) that corresponded to the dodecamer form of the protein $(6n)_2$[91]. The peak at a higher retention volume of ~ 18 mL was estimated to be of 50 kDa (II) corresponding to the dimer form of the protein. This type of distribution of the oligomeric state of the His-tagged *h*Prx3 has been previously described and indicates the stabilization of the dodecamer form by the peptide loop sequence (His-tag)[85]. For the tagged *h*Prx3 protein, the dimer-dodecamer equilibrium seems to be biased, which explains the presence of the large population of dodecamer and a small population of dimers as observed here by SEC-SLS.

For the aromatic muteins (H136F, H136Y, and H136W) the dimer form of the protein was found to dominate with the major peak observed at a retention time of ~18 minutes consistent with the wild-type protein. The alanine mutein H136A in contrast to the aromatic muteins showed a major peak at retention time of ~12 minutes that was consistent with the dodecamer form of the protein and three lesser defined peaks which were estimated to be of 110 kDa, 50 kDa, and 25 kDa.


(A) WT
(6n)2
2n
I
II
Refractive index
Molecular weight (da)
Retention volume (mL)
(B) H136A
Refractive index
Molecular weight (da)
Retention volume (mL)
(C) H136F
Refractive index
Retention volume (mL)

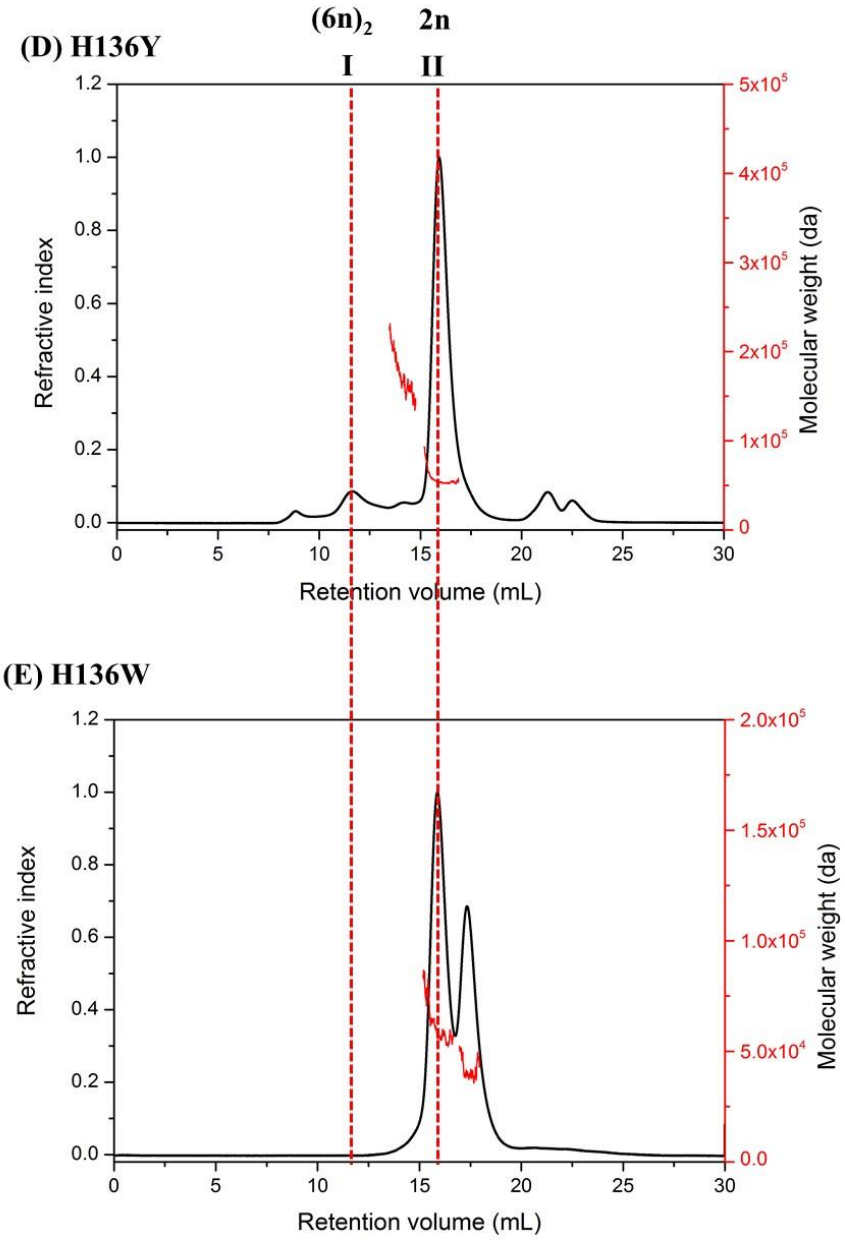


Figure 4.14: *SEC-SLS of hPrx3 and muteins in HEPES, NaCl pH 8.0 on a Superdex 200 10/300 GL column at a flow rate of 0.5 mL/min. The Roman numbers indicate the four prominent protein species separated through chromatography. Dashed lines represent the HMW dodecamer ring form* $(6n)_2$ *and dimer (2n).*

The M_w of the second peak in the analytical SEC-SLS of H136A mutein (Figure 4.14, B) was closest to the theoretical molecular weight of a tetramer as calculated from its amino acid sequence. Table 4.4 lists the M_w calculated for *h*Prx3 WT protein and muteins.

Table 4.4: *List of molecular weights (kDa) obtained for muteins in comparison of hPrx3 WT.*

Protein sample/Oligomers	Molecular weight (kDa)
***h*Prx3 WT**	
Dodecamer	309
Dimer	50
H136A	
Dodecamer	298
Tetramer	115
Dimer	50
Monomer	25
H136Y	
Dodecamer	297
Dimer	50
H136W	
Dimer	50
Monomer	25
H136F	
Dimer	50

The column was pre-equilibrated with Hepes, NaCL pH 8.0 buffer
A Superdex 200 GL 10/300 column was run at a flowrate of 0.5 mL/min. The M_w listed are taken at the midpoint of the eluted peak.

The chromatogram of the protein mutein obtained at a loading concentration of 1 mg/mL was analyzed for the molar mass of the main protein conformation. The molar mass determined was found to be highest at the apex of the peak and fell off on the trailing.

The SEC-SLS provided a good overview of the different oligomeric forms adopted by the *h*Prx3 muteins in solution. The different oligomeric forms identified for the hydrophobic muteins suggested that the specific sequence change of the histidine residue at the homodimer interface affects the monomer-dimer equilibrium that is more biased towards the dimer form.

For the specific mutation of the histidine residue to the large aromatic tryptophan residue, the destabilization of the dimers, however, suggested that the homodimer protein-β interface is weakened. This results in the shift of the monomer-dimer equilibrium towards the monomer form of the protein.

It should be noted that the dodecamer conformation is a stable quaternary assembly of the wild type *h*Prx3 protein dimers adopted in the tagged state of the recombinant protein. Therefore, the oligomeric forms of the *h*Prx3 muteins identified by SEC-SLS consistent in all the four muteins indicated that these oligomeric states are stable assemblies of the tagged recombinant muteins.

4.10.3 Characterization of H136W and H136A muteins

The static light scattering of size excluded protein muteins showed striking results for His-tagged recombinant muteins H136A and H136W. In the case of the aromatic mutein H136W the monomer form of the protein was prefered due to the weakend interactions across the dimer interface. The hydrophobic mutein H136A, however, was found to exist mainly as a dodecamer, though with presence in detectable amounts of the tetramer, dimer and monomer.

To further investigate concentration dependence of the self-assembly of the protein, samples of the muteins at different concentrations were analyzed by SEC-SLS. Figure 4.15, shows the elution profile of H136W mutein at six different concentrations trialed. At lower concentrations, of the mutein at 1 mg/mL and below, the monomer form was more pronounced with a lower proportion of the dimer form. Above 1 mg/mL of mutein in solution, a distinct peak corresponding to the dodecamer form of the protein was present, which became more pronounced with concentration.

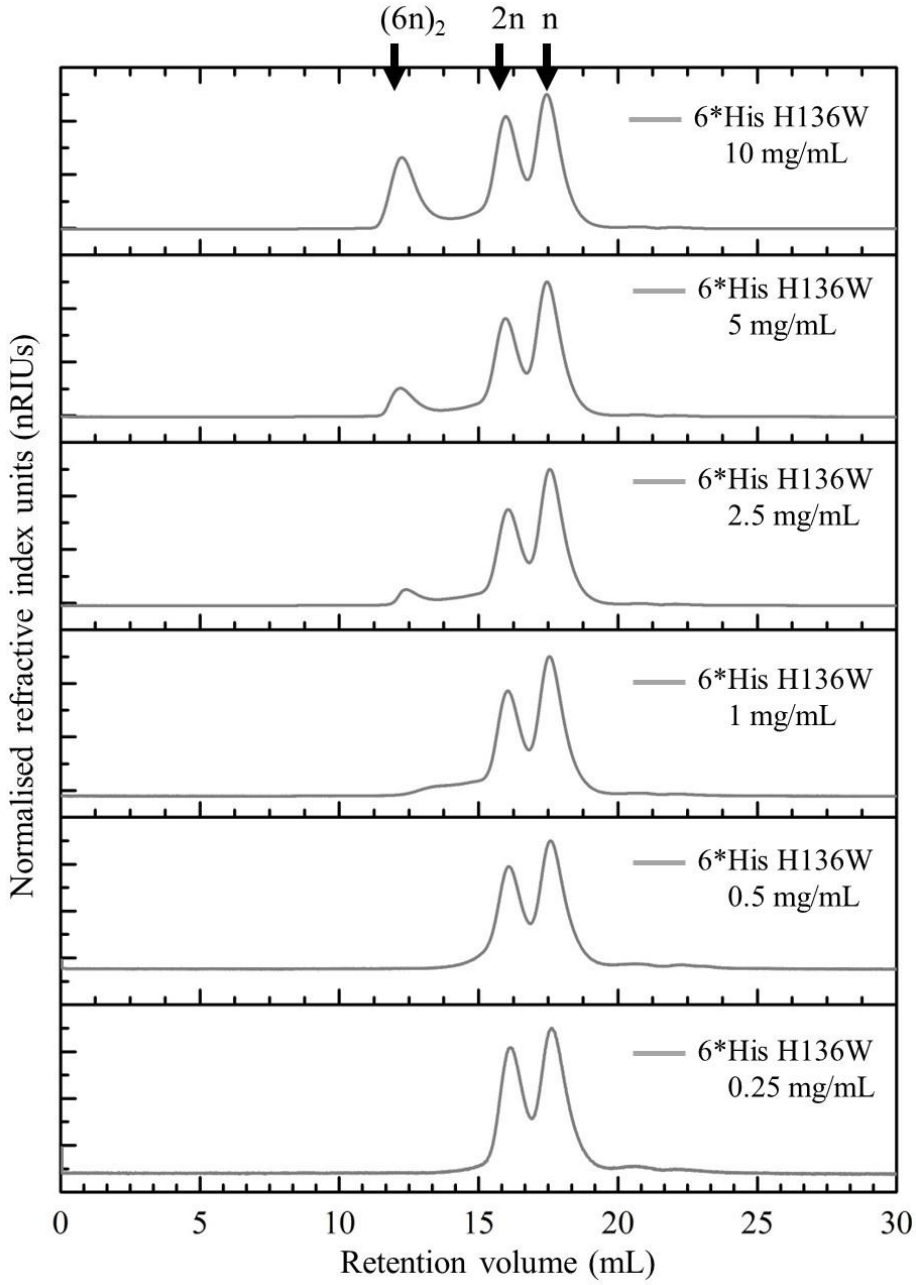


Figure 4.15: *SEC-SLS of H136W 6*His tagged mutein at different concentrations in HEPES, NaCl pH 8.0 buffer. The arrows indicate to the three prominent protein species separated through chromatography, the HMW dodecamer ring form $(6n)_2$, dimer (2n) and monomer (n).*

To study the role of the N-terminal 6*His tag in the assembly of the mutein, a cleaved form of the protein with the tag removed was subjected to SEC-SLS (Figure 4.16).

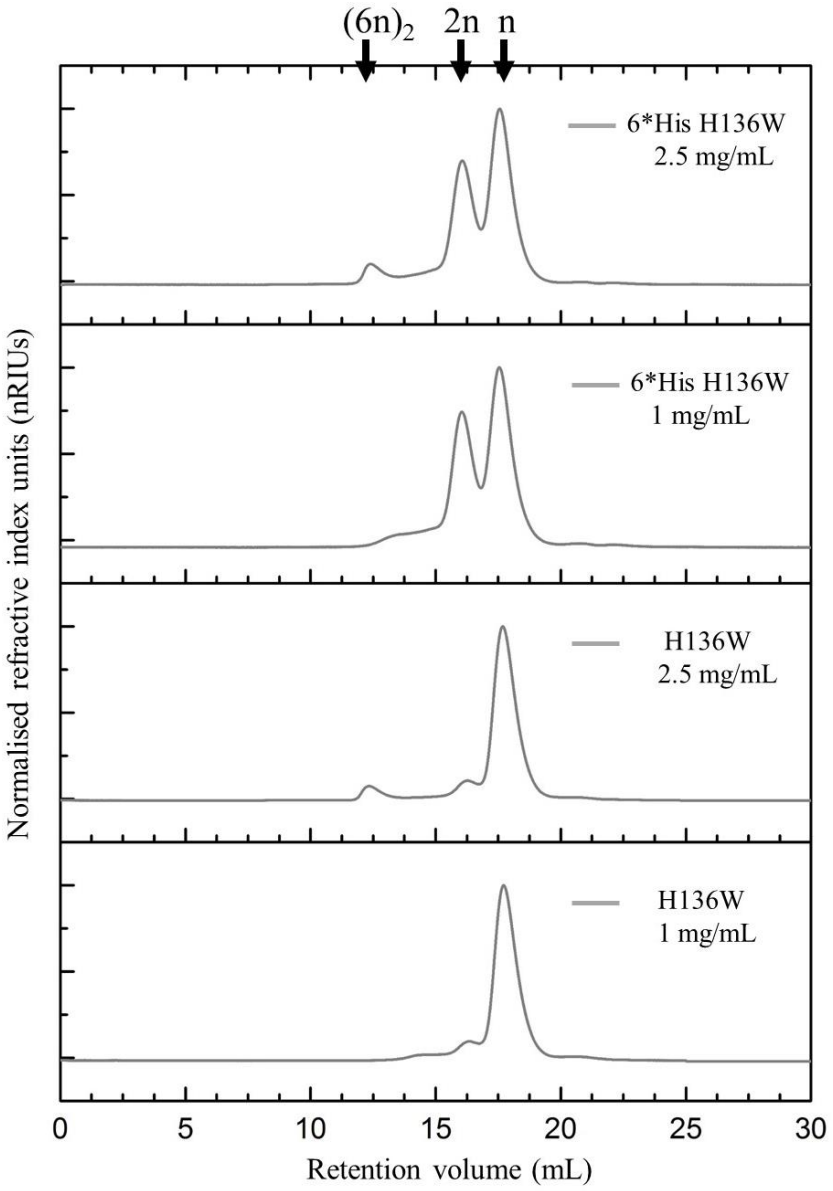


Figure 4.16: *Comparative SEC-SLS of 6*His H136W and H136W (tag cleaved) in HEPES, NaCl pH 8.0 buffer. The arrows indicate to the three prominent protein species separated through chromatography, the HMW dodecamer ring form $(6n)_2$, dimer (2n) and monomer (n).*

Comparison of the SEC-SLS profiles of tagged and untagged H136W muteins showed a striking difference, with the untagged form found to exist mainly as a monomer. This suggested that the specific mutation of the histidine to tryptophan weakens the β-continuous dimer interface. The absence of any substantial rise in the dimer peak but the presence of a dodecamer

peak at a higher concentration of the mutein in solution was also suggestive of a monomer-oligomer equilibrium that is biased towards the dimer. The presence of the dodecamer peak observed for H136W, can be explained by the association of mutein dimers assembled through the protein-β interface at higher concentration into dodecamer toroids. This behavior of the untagged mutein was consistent with the tagged form, where similar monomer to dimer equilibrium existed for different concentration of the protein in solution.

The presence of a relatively large population of H136W dimers observed in the tagged state of the mutein suggested that the monomer-dimer equilibrium was biased, and the tagged state facilitated the association of the protein monomers into dimers. The peptide affinity tag (6*His) exists as a freely accessible N-terminus loop in the 3-dimensional folded form of the protein and near the β9 motif that makes up the β-continuous homodimer interface of *h*Prx3. This raised the possibility of the stabilization of the H136W monomers units into the dimer conformation in the tagged state through interactions arising from the larger surface area of association provided by the peptide tag. The role of the N-terminal tag in the stabilization of the higher order ring-shaped oligomer form of the Prx protein has been previously described [85,92,91].

At dilute concentrations of the cleaved H136W mutein in solution at 1 mg/mL, the monomer form of the protein was pronounced, with a relatively low proportion of the dimer. At twice this concentration of the mutein in solution, the presence of a distinct peak that corresponded to the dodecamer suggested a state of transition of monomers to dimer and the dimers further self-assembling into the higher order dodecamer.

In the case of H136A mutein, the dodecamer form was more pronounced with the relatively lower amount of dimer, and very less monomer peak observed at all concentrations trialed (Figure 4.17).

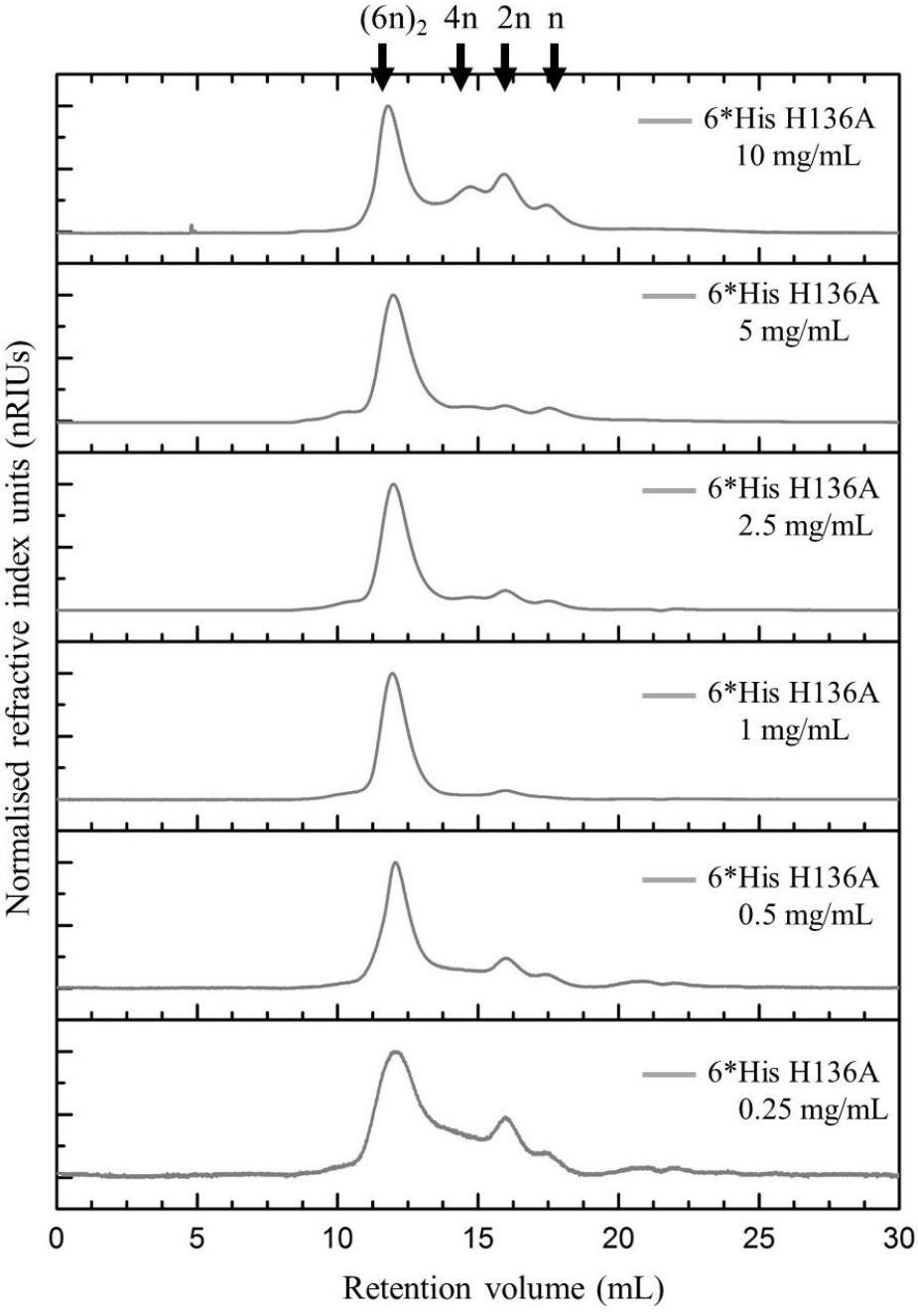


Figure 4.17: *SEC-SLS of H136A 6*His tagged mutein at different concentrations in HEPES, NaCl pH 8.0. The arrows indicate to the four prominent protein species separated through chromatography, the HMW dodecamer ring form $(6n)_2$, possible tetramer (4n), dimer (2n) and monomer (n).*

Above 1 mg/mL in solution, a detectable increase in the dimer peak to that of the monomer suggested a monomer to dimer transition, with a certain ratio of that dimer population self-assembling into dodecamer and a small population of tetramers. This result suggested that the mutation of the histidine to a more hydrophobic alanine residue does not substantially affect the protein homodimer interface.

It is important to note that the dodecamer oligomer form adopted by the *h*Prx3 protein under reducing condition, is facilitated by the partial unfolding of the protein structure and formation of the "Cp-loop"[93]. This partially unfolded structure is in dynamic equilibrium with the fully-folded Prx and the Cp-loop acts as a switch between the dodecamer and dimer. In the mutated form, this dynamic equilibrium seems to be affected, which would explain the distribution of different oligomeric forms of the *h*Prx3 protein detected in the specific case of H136A.

SEC-SLS of untagged H136A mutein showed that most of the protein existed in the higher order dodecamer form, with a relatively low amount of dimer present consistent with that observed for the tagged form (Figure 4.18).

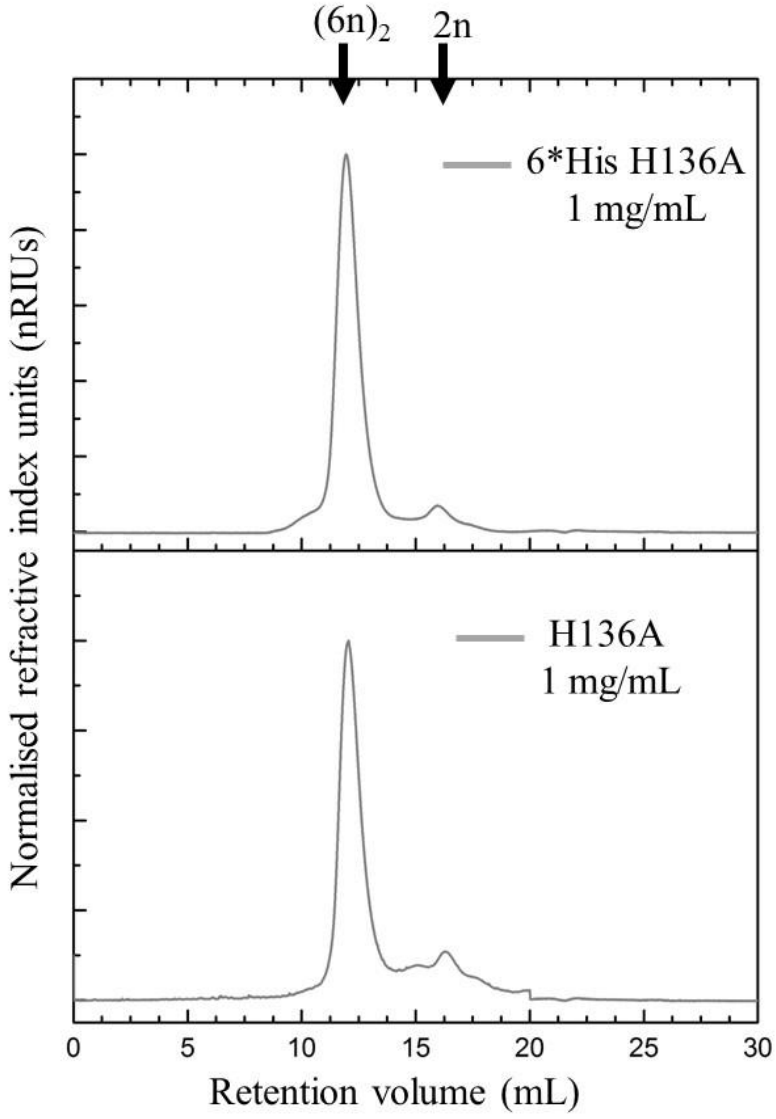


Figure 4.18: *Comparative SEC-SLS of 6*His H136A and H136A (tag cleaved) in HEPES, NaCl pH 8.0 buffer. The arrows indicate the two prominent protein species separated through chromatography, the HMW dodecamer ring form $(6n)_2$ and the dimer (2n).*

Analysis of the tagged and untagged form of the H136W mutein by AUC (Section 2.6.12) at three different concentration confirmed that the monomer form of the protein predominated in solution (Figure 4.19). This suggested that the introduction of the specific hydrophobic mutation in *h*Prx3 protein weakened the of the homodimer protein-β interface.

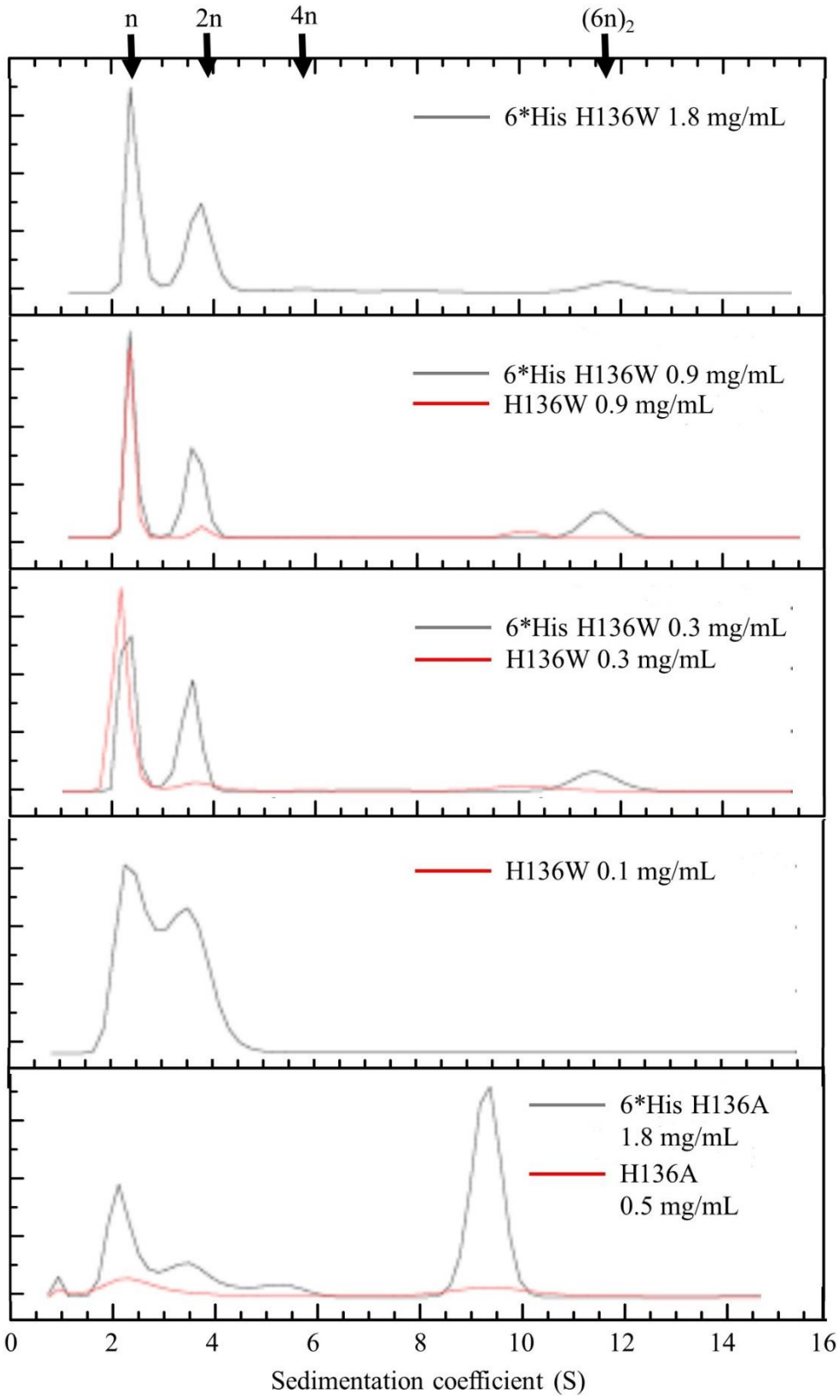


Figure 4.19: *AUC of H136W and H136A muteins in HEPES, NaCl pH 8.0 buffer. The arrows indicate the four prominent protein species sedimenting on centrifugation: n: monomer, 2n: dimer, 4n: tetramer and $(6n)_2$: dodecamer.*

In the case of the H136A mutein, the AUC showed the presence of the different oligomer forms dimer, tetramer and dodecamer, along with some amount of monomer, consistent for the concentrations trailed. The results, thus show that for the mutation of the histidine residue to alanine does not substantially affect the homodimer protein-β interface with the monomer-dimer equilibrium more biased towards the dimer form of the protein. The small population of the H136A mutein monomer form detected could correspond to the small population of non-associated mutein monomers. The AUC also revealed that His-tag supported the formation of protein dimer in H136W mutein, however the tag does not appear to make a substantial difference for H136A mutein (Figure 4.18 and Figure 4.19). It is important to note that the N-terminal region of the protein that contains the His-tag sequence is in close proximity to the homodimer interface in the 3-dimensional fold of the Prx3 protein[18,85]. This raises the possibility of specific side-chain interactions between the His-tag sequence and the dimer interface sequence which stabilizes the otherwise weakened protein-protein interface and favor dimer formation. The role of His-tag in stabilizing protein oligomer conformation has been previously described in Prx3 systems[85,91,92]. The variation of histidine to alanine (H136A) however, does not substantially alter the homodimer interface properties, which explains why no substantial difference is observed in the behavior of the tagged mutein and untagged mutein in solution.

Comparison of the molar masses determined from the apex of the peak data to the theoretical Kd value reported for the wild-type protein generated equilibration constants of dissociation for dimer to monomer in the 2.0 µM to 4.0 µM range for the muteins. The equilibrium constants of dissociation for dimer to monomer was judged to be > 2-4 µM as complete dissociation of the dimer to monomer was not observed for the lowest concentration trailed in this work.

4.11 Discussion

The histidine residue in the isolated Prx3 peptide sequence was chosen to be varied as it was found to be highly conserved within the parent β-continuous homodimer interface in Prx3 protein and within hydrogen bonding distance to neighboring serine and asparagine residues. Peptide variants designed with the histidine replaced with hydrophobic residues all formed birefringent hydrogels with liquid crystalline textures observed for all the four peptide acetates in water. The variant peptide sequences showed a propensity to exist as β-sheets at dilute concentration in solution[36] similar to the wild-type peptide. This indicated that the specific variation of the histidine residue in the peptide does not alter backbone hydrogen bonding and thus β-sheet conformation is retained.

ThT binding assays showed that all the variant peptide acetates in water associated with the amyloid-specific ThT dye in a reversible manner, though to a lower extent, but consistent with the parent peptide. The binding assays further showed that the Prx3 peptide sequence with a specific variation of the histidine to the aromatic tyrosine or tryptophan has higher critical assembly concentration. Peptide self-assembly studies as a function of time showed that the specific aromatic variations of the Prx3 peptide (H3W and H3Y) exhibited a two-state model of self-assembly consistent with the presence of a lag phase. Time course studies on amyloid fibril formation by proteins and peptides usually exhibit a lag or nucleation phase and growth phase which occurs rapidly with the addition of monomers and stationary phase that is predominated by mature fibrils[94,95,96,97]. The rate of self-assembly can be changed by the addition the nuclei or seeds that shorten the lag phase. The absence of any substantial lag phase at a higher concentration of the peptide acetate in water can be explained by the solute effects on the monomer-aggregate equilibrium.

Apparent rate constants of aggregation revealed that the specific change in the sequence from histidine to phenylalanine resulted in faster assembly kinetics. The specific change in the peptide sequence to the more hydrophobic phenylalanine residue results in burying of the phenylalanine side-chains away from solvent molecules and the specific aromatic interactions between phenylalanine side-chains favour faster aggregation kinetics. Phenylalanine has been noted to be an important residue in driving amyloid assembly and frequently encountered in varied in amyloidogenic motifs reported to date[50,51]. Furthermore, peptides comprised of only two repeating phenylalanine residues have been shown to self-assemble into full length amyloid fibrils largely driven by aromatic interactions[12,13]. Diphenylalanine motifs are shortest self-assembling peptide motifs developed using the "reductionist approach" by Gazit & co-workers[50,51] and have been successfully used as "molds" for fabrication of metal nanowires. Furthermore, self-assembling diphenylalanine peptides can be grown on surfaces as vertical nanowires that act as attachment surface for 3D cell culture for development of biosensing devices[14]. Thus, peptide sequences containing phenylalanine residue provide a good starting point to build peptide based building blocks that can self-assemble into discrete nanostructures to be used a biomaterials.

Thermodynamic studies on pressure and solute effects on the self-assembly of biomolecules in solution show that the association-dissociation equilibrium is affected by the change in the environment including pressure, temperature and number of solute molecules [98,38,99,39]. Also, it has been noted that lyophilized peptide acetate sample shows the presence of amyloid nucleates[97]. This explains the substantial ThT fluorescence observed immediately after dilution of peptide acetate powder in water and for concentrated peptide acetate in water samples. The rise in the ThT fluorescence observed over time is consistent with the growth of the peptide fibrillar structures with the addition of peptide monomers[96,97].

TEM and SAXS corroborated evidence that all the designed variant peptide sequences self-assembled into β-sheet based nanostructures with extended morphologies consistent with those described for the parent Prx3 peptide. SAXS patterns of the peptide sequences showed that the hydrophobic peptide variants have a similar β-sheet organization at the molecular level which is repetitive over long distances, consistent with the parent peptide. The molecular organization observed for the variant peptide sequences are similar to that observed for different peptide sequences that self-assembled into amyloid-like fibrils[66,67,69,100]. Some differences observed at the small angles can be correlated to the slight differences observed by TEM concerning the symmetries which is attributed to the different lamination of the β-sheets in the extended nanostructures. This difference in the symmetry of the β-sheets can be associated with the different morphologies observed for the variant peptides and furthermore to the polymorphism observed in the nanotapes formed by the variant peptide H3Y[101,102]. FTIR further confirmed that the backbone conformation of both wild-type Prx3 peptide and the hydrophobic variants was that of an antiparallel β-sheet type with the characteristic combinative vibrations[77,78].

A mutational approach used to explore the effect of variation of the histidine residue at the parent Prx3 homodimer interface changed the high order dodecamer assembly of the protein. However, the dimer form of the Prx3 protein was found to be largely stable across all the hydrophobic muteins studied with the only exception being the H136W mutein. The change in the monomer-dimer equilibrium observed in the case of *h*Prx3 mutein H136W can be attributed to the replacement of the imidazole side chain group by the bigger aromatic side chain. It is important to note that the β9-β9' motif within the homodimer interface of the *h*Prx3 dodecamer has restricted movement. Thus, the weakening of the interface can be attributed to the localized steric hindrance caused due to the introduction of the tryptophan residue having a large aromatic side chain group. This, eventually destabilizes the interface and can explain the

predominance of monomer species observed in solution for the H136W mutein. The behavior of the parent mutein was found to be in contrast to the corresponding isolated peptide variant sequence H3W that self-assembled into extended β-sheet based nanostructures. This suggested that the isolated peptide sequence interacts in a different way to the parent interface. In the isolated form, the smaller size and length of the sequence offers a higher degree of flexibility and freedom to associate into β-sheets, overcoming the steric hindrance due to the large aromatic group. Within the 3-dimensional fold of the Prx3 protein this flexibility of the β-motif is restricted, and the steric effects due to the large aromatic tryptophan residue limit the association of the protein monomers.

Overall, the results showed that the peptide sequence corresponding to the homodimer protein-β interface of Prx3 protein is stable to the specific hydrophobic variation of the histidine residue. Of the different peptide sequences studied the H3F peptide was the most aggregation prone. The different muteins studied designed with the corresponding variation of the histidine residue all behaved similarly except the H136W where the monomer form was preferred. The results thus support the variant peptide sequences as molecular tectons with self-assembling properties stemming from the parent protein-β interface. The results shown here augur well with those observed with proteins and peptide sequences that self-assemble into amyloid-like fibrils and showcase peptide sequence derived from protein interface as stable molecular tectons. Such stable self-assembling peptide tectons show potential to be used as building blocks for the design of tailored biomaterials for fabrication of specific patterns through a bottom-up approach.

4.12 References

1. Jun, S., Hong, Y., Imamura, H., Ha, B. Y., Bechhoefer, J. & Chen, P. (2004). Self-Assembly of the Ionic Peptide EAK16: The effect of charge distributions on self-assembly. *Biophysical Journal* **87**, 1249-1259.
2. Jeon, J. & Shell, M S. (2012). Charge effects on the fibril-forming peptide KTVIIE: a two-dimensional replica exchange simulation study. *Biophysical Journal* **102**, 1952-1960.
3. Zhao, X., Pan, F. & Lu, J. R. (2008). Recent development of peptide self-assembly. *Progress in Natural Science* **18**, 653-660.
4. Lee, N. R., Bowerman, C. J. & Nilsson, B. L. (2013). Effects of varied sequence pattern on the self-assembly of amphipathic peptides. *Biomacromolecules* **14**, 3267-3277.
5. Caplan, M. R., Moore, P. N., Zhang, S., Kamm, R. D. & Lauffenburger, D. A. (2000). Self-assembly of a β-sheet protein governed by relief of electrostatic repulsion relative to van der Waals attraction. *Biomacromolecules* **1**, 627-631.
6. Caplan, M. R., Schwartzfarb, E. M., Zhang, S., Kamm, R. D. & Lauffenburger, D. A. (2002). Control of self-assembling oligopeptide matrix formation through systematic variation of amino acid sequence. *Biomaterials* **23**, 219-227.
7. López de la Paz, M., Goldie, K., Zurdo, J., Lacroix, E., Dobson, C. M., Hoenger, A. & Serrano, L. (2002). De novo designed peptide-based amyloid fibrils. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 16052-16057.
8. Hamley, I. W. (2007). Peptide fibrillization. *Angewandte Chemie International Edition in English* **46**, 8128-8147.
9. Gazit, E. (2005). Mechanisms of amyloid fibril self-assembly and inhibition. model short peptides as a key research tool. *FEBS Journal* **272**, 5971-5978.
10. Gazit, E. (2007). Self assembly of short aromatic peptides into amyloid fibrils and related nanostructures. *Prion* **1**, 32-35.
11. Reches, M. & Gazit, E. (2004). Amyloidogenic hexapeptide fragment of medin: homology to functional islet amyloid polypeptide fragments. *Amyloid* **11**, 81-89.
12. Tamamis, P., Adler-Abramovich, L., Reches, M., Marshall, K., Sikorski, P., Serpell, L., Gazit, E. & Archontis, G. (2009). Self-assembly of phenylalanine oligopeptides: insights from experiments and simulations. *Biophysical Journal* **96**, 5020-5029.

13. Maeda, I., Taniguchi, S., Watanabe, N., Inoue, A., Yamasaki, Y. & Nose, T. (2015). Design of phenylalanine-containing elastin-derived peptides exhibiting highly potent self-assembling capability. *Protein and Peptide Letters* **22**, 934-939.
14. Sasso, L., Vedarethinam, I., Emneus, J., Svendsen, W. E. & Castillo-Leon, J. (2012). Self-assembled diphenylalanine nanowires for cellular studies and sensor applications. *Journal of Nanoscience and Nanotechnology* **12**, 3077-3083.
15. Valéry, C., Pouget, E., Pandit, A., Verbavatz, J.-M., Bordes, L., Boisdé, I., Cherif-Cheikh, R., Artzner, F. & Paternostre, M. (2008). Molecular origin of the self-assembly of lanreotide into nanotubes: a mutational approach. *Biophysical Journal* **94**, 1782-1795.
16. Armstrong, A. H., Chen, J., McKoy, A. F. & Hecht, M. H. (2011). Mutations that replace aromatic side chains promote aggregation of the Alzheimer's Aβ peptide. *Biochemistry* **50**, 4058-4067.
17. Valery, C., Pandey, R. & Gerrard, J. A. (2013). Protein β-interfaces as a generic source of native peptide tectons. *Chemical Communications* **49**, 2825-2827.
18. Cao, Z., Roszak, A. W., Gourlay, L. J., Lindsay, J. G. & Isaacs, N. W. (2005). Bovine mitochondrial peroxiredoxin III forms a two-ring catenane. *Structure* **13**, 1661-1664.
19. Hamley, I. W., Kirkham, S., Dehsorkhi, A., Castelletto, V., Adamcik, J., Mezzenga, R., Ruokolainen, J., Mazzuca, C., Gatto, E., Venanzi, M., Placidi, E., Bilalis, P. & Iatrou, H. (2014). Self-assembly of a model peptide incorporating a hexa-histidine sequence attached to an oligo-alanine sequence, and binding to gold NTA/nickel nanoparticles. *Biomacromolecules* **15**, 3412-3420.
20. Qin, Z.-l., Zheng, Y. & Kielian, M. (2009). Role of conserved histidine residues in the low-pH dependence of the semliki forest virus fusion protein. *Journal of Virology* **83**, 4670-4677.
21. Tu, Z., Young, A., Murphy, C. & Liang, J. F. (2009). The pH sensitivity of histidine-containing lytic peptides. *Journal of peptide science* **15**, 790-795.
22. Kalani, Mohamad R., Moradi, A., Moradi, M. & Tajkhorshid, E. (2013). Characterizing a histidine switch controlling pH-dependent conformational changes of the influenza virus hemagglutinin. *Biophysical Journal* **105**, 993-1003.
23. Mair, C. M., Meyer, T., Schneider, K., Huang, Q., Veit, M. & Herrmann, A. (2014). A histidine residue of the influenza virus hemagglutinin controls the pH dependence of the conformational change mediating membrane fusion. *Journal of Virology* **88**, 13189-13200.

24. Carpena, X., Vidossich, P., Schroettner, K., Calisto, B. M., Banerjee, S., Stampler, J., Soudi, M., Furtmuller, P. G., Rovira, C., Fita, I. & Obinger, C. (2009). Essential role of proximal histidine-asparagine interaction in mammalian peroxidases. *Journal of Biological Chemistry* **284**, 25929-25937.
25. Passardi, F., Bakalovic, N., Teixeira, F. K., Margis-Pinheiro, M., Penel, C. & Dunand, C. (2007). Prokaryotic origins of the non-animal peroxidase superfamily and organelle-mediated transmission to eukaryotes. *Genomics* **89**, 567-479.
26. Birktoft, J. J. & Banaszak, L. J. (1983). The presence of a histidine-aspartic acid pair in the active site of 2-hydroxyacid dehydrogenases: X-ray refinement of cytoplasmic malate dehydrogenase. *Journal of Biological Chemistry* **258**, 472-482.
27. Valéry, C., Deville-Foillard, S., Lefebvre, C., Taberner, N., Legrand, P., Meneau, F., Meriadec, C., Delvaux, C., Bizien, T., Kasotakis, E., Lopez-Iglesias, C., Gall, A., Bressanelli, S., Le Du, M.-H., Paternostre, M. & Artzner, F. (2015). Atomic view of the histidine environment stabilizing higher-pH conformations of pH-dependent proteins. *Nature Communications* **6**, 7771-7778.
28. Kirkwood, J. G. & Buff, F. P. (1951). The statistical mechanical theory of solutions. *Journal of Chemical Physics* **19**, 774-777.
29. Ben-Naim, A. (2006). Molecular theory of solutions. *Oxford University Press*.
30. Ben-Naim, A. (2012). Theoretical aspects of pressure and solute denaturation of proteins: A Kirkwood-buff-theory approach. *Journal of Chemical Physics* **137**, 235102-235108.
31. Ben-Naim, A. (2011). Molecular theory of water and aqueous solutions: The role of water in protein folding, self-assembly and molecular recognition. World Scientific.
32. Ben-Naim, A. (2006). On the driving forces for protein-protein association. *Journal of Chemical Physics* **125**, 024901-024910.
33. Jarrett, J. T. & Lansbury, P. T., Jr. (1993). Seeding "one-dimensional crystallization" of amyloid: a pathogenic mechanism in Alzheimer's disease and scrapie? *Cell* **73**, 1055-1058.
34. Bolisetty, S., Adamcik, J. & Mezzenga, R. (2011). Snapshots of fibrillation and aggregation kinetics in multistranded amyloid β-lactoglobulin fibrils. *Soft Matter* **7**, 493-499.

35. Pomerantz, W. C., Yuwono, V. M., Pizzey, C. L., Hartgerink, J. D., Abbott, N. L. & Gellman, S. H. (2008). Nanofibers and lyotropic liquid crystals from a class of self-assembling β-peptides. *Angewandte Chemie International Edition in English* **47**, 1241-1244.

36. Greenfield, N. J. (2006). Using circular dichroism spectra to estimate protein secondary structure. *Nature Protocols* **1**, 2876-2890.

37. Knapman, T. W., Aggeli, A. & Ashcroft, A. E. (2008). Critical concentrations of β-sheet peptide self-assembly quantified directly by nanoelectrospray ionization mass spectrometry. *Rapid Communications in Mass Spectrometry* **22**, 1611-1614.

38. Pierce, V., Kang, M., Aburi, M., Weerasinghe, S. & Smith, P. E. (2008). Recent Applications of Kirkwood–Buff theory to biological systems. *Cell Biochemistry and Biophysics* **50**, 1-22.

39. Gee, M. B. & Smith, P. E. (2009). Kirkwood–Buff theory of molecular and protein association, aggregation, and cellular crowding. *Journal of Chemical Physics* **131**, 165101-165111.

40. Campioni, S., Mossuto, M. F., Torrassa, S., Calloni, G., de Laureto, P. P., Relini, A., Fontana, A. & Chiti, F. (2008). Conformational properties of the aggregation precursor state of HypF-N. *Journal of Molecular Biology* **379**, 554-567.

41. Kodali, R. & Wetzel, R. (2007). Polymorphism in the intermediates and products of amyloid assembly. *Current Opinion in Structural Biology* **17**, 48-57.

42. Otzen, D. E. (2013). Amyloid fibrils and prefibrillar aggregates: molecular and biological properties. *John Wiley & sons*.

43. Xue, W. F., Homans, S. W. & Radford, S. E. (2008). Systematic analysis of nucleation-dependent polymerization reveals new insights into the mechanism of amyloid self-assembly. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 8926-31.

44. Streets, A. M., Sourigues, Y., Kopito, R. R., Melki, R. & Quake, S. R. (2013). Simultaneous measurement of amyloid fibril formation by dynamic light scattering and fluorescence reveals complex aggregation kinetics. *PLoS One* **8**, 1-10.

45. Chiti, F., Taddei, N., Baroni, F., Capanni, C., Stefani, M., Ramponi, G. & Dobson, C. M. (2002). Kinetic partitioning of protein folding and aggregation. *Nature Structural and Molecular Biology* **9**, 137-143.

46. Hortschansky, P., Schroeckh, V., Christopeit, T., Zandomeneghi, G. & Fändrich, M. (2005). The aggregation kinetics of Alzheimer's β-amyloid peptide is controlled by stochastic nucleation. *Protein Science* **14**, 1753-1759.

47. Calamai, M., Taddei, N., Stefani, M., Ramponi, G. & Chiti, F. (2003). Relative influence of hydrophobicity and net charge in the aggregation of two homologous proteins. *Biochemistry* **42**, 15078-15083.

48. Chiti, F., Stefani, M., Taddei, N., Ramponi, G. & Dobson, C. M. (2003). Rationalization of the effects of mutations on peptide and protein aggregation rates. *Nature* **424**, 805-808.

49. DuBay, K. F., Pawar, A. P., Chiti, F., Zurdo, J., Dobson, C. M. & Vendruscolo, M. (2004). Prediction of the absolute aggregation rates of amyloidogenic polypeptide chains. *Journal of Molecular Biology* **341**, 1317-1326.

50. Azriel, R. & Gazit, E. (2001). Analysis of the minimal amyloid-forming fragment of the islet amyloid polypeptide. An experimental support for the key role of the phenylalanine residue in amyloid formation. *Journal of Biological Chemistry* **276**, 34156-34161.

51. Gazit, E. (2007). Self assembly of short aromatic peptides into amyloid fibrils and related nanostructures. *Prion* **1**, 32-35.

52. Makin, O. S., Atkins, E., Sikorski, P., Johansson, J. & Serpell, L. C. (2005). Molecular basis for amyloid fibril formation and stability. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 315-320.

53. Chelli, R., Gervasio, F. L., Procacci, P. & Schettino, V. (2002). Stacking and T-shape competition in aromatic-aromatic amino acid interactions. *Journal of the American Chemical Society* **124**, 6133-43.

54. Javid, N., Roy, S., Zelzer, M., Yang, Z., Sefcik, J. & Ulijn, R. V. (2013). Cooperative self-assembly of peptide gelators and proteins. *Biomacromolecules* **14**, 4368-4376.

55. van der Zwaag, D., Pieters, P. A., Korevaar, P. A., Markvoort, A. J., Spiering, A. J. H., de Greef, T. F. A. & Meijer, E. W. (2015). Kinetic analysis as a tool to distinguish pathway complexity in molecular assembly: an unexpected outcome of structures in competition. *Journal of the American Chemical Society* **137**, 12677-12688.

56. Petkova, A. T., Leapman, R. D., Guo, Z., Yau, W. M., Mattson, M. P. & Tycko, R. (2005). Self-propagating, molecular-level polymorphism in Alzheimer's β-amyloid fibrils. *Science* **307**, 262-265.

57. Verel, R., Tomka, I. T., Bertozzi, C., Cadalbert, R., Kammerer, R. A., Steinmetz, M. O. & Meier, B. H. (2008). Polymorphism in an Amyloid-Like Fibril-Forming Model Peptide. *Angewandte Chemie International Edition in English* **47**, 5842-5845.

58. Fitzpatrick, A. W. P., Debelouchina, G. T., Bayro, M. J., Clare, D. K., Caporini, M. A., Bajaj, V. S., Jaroniec, C. P., Wang, L., Ladizhansky, V., Müller, S. A., MacPhee, C. E., Waudby, C. A., Mott, H. R., De Simone, A., Knowles, T. P. J., Saibil, H. R., Vendruscolo, M., Orlova, E. V., Griffin, R. G. & Dobson, C. M. (2013). Atomic structure and hierarchical assembly of a cross-β amyloid fibril. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 5468-5473.

59. Krasnoslobodtsev, A. V., Deckert-Gaudig, T., Zhang, Y., Deckert, V. & Lyubchenko, Y. L. (2016). Polymorphism of amyloid fibrils formed by a peptide from the yeast prion protein Sup35: AFM and tip-enhanced raman scattering studies. *Ultramicroscopy* **165**, 26-33.

60. Paravastu, A. K., Petkova, A. T. & Tycko, R. (2006). Polymorphic fibril formation by residues 10–40 of the Alzheimer's β-amyloid peptide. *Biophysical Journal* **90**, 4618-4629.

61. van der Wel, P. C. A., Lewandowski, J. R. & Griffin, R. G. (2007). Solid-state NMR study of amyloid nanocrystals and fibrils formed by the peptide GNNQQNY from yeast prion protein Sup35p. *Journal of the American Chemical Society* **129**, 5117-5130.

62. Antzutkin, O. N., Leapman, R. D., Balbach, J. J. & Tycko, R. (2002). Supramolecular structural constraints on Alzheimer's β-amyloid fibrils from electron microscopy and solid-state NMR. *Biochemistry* **41**, 15436-15450.

63. Petkova, A. T., Buntkowsky, G., Dyda, F., Leapman, R. D., Yau, W. M. & Tycko, R. (2004). Solid state NMR reveals a pH-dependent antiparallel β-Sheet registry in fibrils formed by a β-Amyloid peptide. *Journal of Molecular Biology* **335**, 247-260.

64. Fitzpatrick, A. W., Knowles, T. P., Waudby, C. A., Vendruscolo, M. & Dobson, C. M. (2011). Inversion of the balance between hydrophobic and hydrogen bonding interactions in protein folding and aggregation. *PLoS Computional Biology* **7**, 1-10.

65. Segman, S., Lee, M.-r., Vaiser, V., Gellman, S. H. & Rapaport, H. (2010). Highly stable pleated-sheet secondary structure in assemblies of amphiphilic α/β-peptides at the air–water interface. *Angewandte Chemie International Edition in English* **49**, 716-719.

66. Serpell, L. C. (2000). Alzheimer's amyloid fibrils: structure and assembly. *Biochimica et Biophysica Acta: Molecular Basis of Disease* **1502**, 16-30.

67. Sunde, M., Serpell, L. C., Bartlam, M., Fraser, P. E., Pepys, M. B. & Blake, C. C. (1997). Common core structure of amyloid fibrils by synchrotron X-ray diffraction. *Journal of Molecular Biology* **273**, 729-739.

68. Guinier, A. (1994). X-ray diffraction in crystals, imperfect crystals, and amorphous bodies. *Dover*.

69. van Grondelle, W., Iglesias, C. L., Coll, E., Artzner, F., Paternostre, M., Lacombe, F., Cardus, M., Martinez, G., Montes, M., Cherif-Cheikh, R. & Valery, C. (2007). Spontaneous fibrillation of the native neuropeptide hormone Somatostatin-14. *Journal of Structural Biology* **160**, 211-223.

70. Haris, P. I. & Chapman, D. (1995). The conformational analysis of peptides using Fourier transform infrared spectroscopy. *Biopolymers* **37**, 251-263.

71. Susi, H. & Byler, D. M. (1987). Fourier transform infrared study of proteins with parallel β-chains. *Archives of Biochemistry and Biophysics* **258**, 465-469.

72. Stuart, B. H. & Ando, D. J. (1997). Biological applications of infrared spectroscopy. *John Wiley & sons*.

73. Krimm, S. & Bandekar, J. (1986). Vibrational spectroscopy and conformation of peptides, polypeptides, and proteins. *Advances in Protein Chemistry* **38**, 181-364.

74. Bandekar, J. (1992). Amide modes and protein conformation. *Biochimica et Biophysica Acta* **1120**, 123-143.

75. Goormaghtigh, E., Cabiaux, V. & Ruysschaert, J.-M. (1994). Determination of Soluble and Membrane Protein Structure by Fourier Transform Infrared Spectroscopy. In Physicochemical Methods in the Study of Biomembranes. *Springer science* **23**, 329-362.

76. Pelton, J. T. & McLean, L. R. (2000). Spectroscopic methods for analysis of protein secondary structure. *Analytical Biochemistry* **277**, 167-176.

77. Zandomeneghi, G., Krebs, M. R. H., McCammon, M. G. & Fändrich, M. (2004). FTIR reveals structural differences between native β-sheet proteins and amyloid fibrils. *Protein Science* **13**, 3314-3321.

78. Barth, A. (2007). Infrared spectroscopy of proteins. *Biochimica et Biophysica Acta: Bioenergetics* **1767**, 1073-1101.

79. van Grondelle, W., Lecomte, S., Lopez-Iglesias, C., Manero, J.-M., Cherif-Cheikh, R., Paternostre, M. & Valery, C. (2013). Lamination and spherulite-like compaction of a hormone's native amyloid-like nanofibrils: spectroscopic insights into key interactions. *Faraday Discussions* **166**, 163-180.

80. Stroud, J. C., Liu, C., Teng, P. K. & Eisenberg, D. (2012). Toxic fibrillar oligomers of amyloid-β have cross-β structure. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 7717-7722.

81. Valery, C., Artzner, F., Robert, B., Gulick, T., Keller, G., Grabielle-Madelmont, C., Torres, M. L., Cherif-Cheikh, R. & Paternostre, M. (2004). Self-association process of a peptide in solution: from β-sheet filaments to large embedded nanotubes. *Biophysical Journal* **86**, 2484-501.

82. Kawasaki, T., Fujioka, J., Imai, T., Torigoe, K. & Tsukiyama, K. (2014). Mid-infrared free-electron laser tuned to the amide I band for converting insoluble amyloid-like protein fibrils into the soluble monomeric form. *Lasers in Medical Science* **29**, 1701-1707.

83. Kawasaki, T., Yaji, T., Imai, T., Ohta, T. & Tsukiyama, K. (2014). Synchrotron-infrared microscopy analysis of amyloid fibrils irradiated by mid-infrared free-electron laser. *American Journal of Analytical Chemistry* **5**, 384-394.

84. Hall, A., Karplus, P. A. & Poole, L. B. (2009). Typical 2-Cys peroxiredoxins: structures, mechanisms and functions. *The FEBS Journal* **276**, 2469-2477.

85. Barranco-Medina, S., Lázaro, J.-J. & Dietz, K.-J. (2009). The oligomeric conformation of peroxiredoxins links redox state to function. *FEBS Letters* **583**, 1809-1816.

86. Qiu, W., Dong, A., Pizarro, J. C., Botchkarsev, A., Min, J., Wernimont, A. K., Hills, T., Hui, R. & Artz, J. D. (2012). Crystal structures from the plasmodium peroxiredoxins: new insights into oligomerization and product binding. *BMC Structural Biology* **12**, 1-18.

87. Saccoccia, F., Di Micco, P., Boumis, G., Brunori, M., Koutris, I., Miele, A. E., Morea, V., Sriratana, P., Williams, D. L., Bellelli, A. & Angelucci, F. (2012). Moonlighting by different stressors: crystal structure of the chaperone species of a 2-Cys peroxiredoxin. *Structure* **20**, 429-439.

88. Matthes, D., Gapsys, V. & de Groot, B. L. (2012). Driving forces and structural determinants of steric zipper peptide oligomer formation elucidated by atomistic simulations. *Journal of Molecular Biology* **421**, 390-416.

89. Rosano, G. L. & Ceccarelli, E. A. (2014). Recombinant protein expression in *E. coli*: Advances and challenges. *Frontiers in Microbiology* **5**, 1-17.

90. Gazit, E. (2002). A possible role for pi-stacking in the self-assembly of amyloid fibrils. *FASEB Journal* **16**, 77-83.

91. Phillips, A. J., Littlejohn, J., Yewdall, N. A., Zhu, T., Valéry, C., Pearce, F. G., Mitra, A. K., Radjainia, M. & Gerrard, J. A. (2014). Peroxiredoxin is a versatile self-assembling tecton for protein nanotechnology. *Biomacromolecules* **15**, 1871-1881.

92. Gretes, M. C. & Karplus, P. A. (2013). Observed octameric assembly of a plasmodium yoelii peroxiredoxin can be explained by the replacement of native "ball-and-socket" interacting residues by an affinity tag. *Protein Science* **22**, 1445-1452.

93. Wood, Z. A., Poole, L. B., Hantgan, R. R. & Karplus, P. A. (2002). Dimers to doughnuts: redox-sensitive oligomerization of 2-cysteine peroxiredoxins. *Biochemistry* **41**, 5493-504.

94. Naiki, H., Hashimoto, N., Suzuki, S., Kimura, H., Nakakuki, K. & Gejyo, F. (1997). Establishment of a kinetic model of dialysis-related amyloid fibril extension *in vitro*. *Amyloid* **4**, 223-232.

95. Serio, T. R., Cashikar, A. G., Kowal, A. S., Sawicki, G. J., Moslehi, J. J., Serpell, L., Arnsdorf, M. F. & Lindquist, S. L. (2000). Nucleated conformational conversion and the replication of conformational information by a prion determinant. *Science* **289**, 1317-1321.

96. Lee, J., Culyba, E. K., Powers, E. T. & Kelly, J. W. (2011). Amyloid-β forms fibrils by nucleated conformational conversion of oligomers. *Nature Chemical Biology* **7**, 602-609.

97. Naldi, M., Fiori, J., Pistolozzi, M., Drake, A. F., Bertucci, C., Wu, R., Mlynarczyk, K., Filipek, S., De Simone, A. & Andrisano, V. (2012). Amyloid β-peptide 25-35 self-assembly and its inhibition: a model undecapeptide system to gain atomistic and secondary structure details of the Alzheimer's disease process and treatment. *ACS Chemical Neuroscience* **3**, 952-962.

98. Jiao, Y. & Smith, P. E. (2011). Fluctuation theory of molecular association and conformational equilibria. *Journal of Chemical Physics* **135**, 014502-014513.

99. Stites, W. E. (1997). Protein-protein interactions: Interface structure, binding thermodynamics, and mutational analysis. *Chemical Reviews* **97**, 1233-1250.

100. Langkilde, A. E., Morris, K. L., Serpell, L. C., Svergun, D. I. & Vestergaard, B. (2015). The architecture of amyloid-like peptide fibrils revealed by X-ray scattering, diffraction and electron microscopy. *Acta Crystallography Section D: Biological Crystallography* **71**, 882-895.

101. Colletier, J.-P., Laganowsky, A., Landau, M., Zhao, M., Soriaga, A. B., Goldschmidt, L., Flot, D., Cascio, D., Sawaya, M. R. & Eisenberg, D. (2011). Molecular basis for amyloid-β polymorphism. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 16938-16943.

102. Tycko, R. & Wickner, R. B. (2013). Molecular structures of amyloid and prion Fibrils: Consensus vs. Controversy. *Accounts of Chemical Research* **46**, 1487-1496.1.

5 Chapter Five

Templating higher order nanotubular assembly of *h*Prx3 protein

5.1 Introduction

Orchestrating the assembly of biological building blocks into hierarchical structures that can be used for fabrication of specific patterns requires the design of stable building blocks that can be triggered to self-assemble. Molecular self-assembly provides a versatile pathway for production of nanoscale materials that can be used in bionanotechnology applications such as templating minerals for production of conductive materials and fabrication of nanodevices. Exploiting the chemistry of biological tectons through a bottom-up approach has been used to create well-ordered structures in the nano- and micro-scale dimension with specific properties that can be used to trigger self-assembly. The use of simple biological molecules as scaffolds and templates has allowed for the production of a large number of varied yet customized non-biological structures. Important examples include DNA origami[1,2] and RNA origami[3,4] that have been successfully used to create a range of different nanoscale shapes and patterns.

Proteins offer some major advantages over nucleic acids as biological tectons since the range of different 3-dimensional structures possible is vast and they offer a wide range of surface chemistry that can be tuned for functionalization[5,6,7]. However, such templating remains challenging due to a lack of stable tectons that can direct such assembly. Naturally occurring proteins can organize themselves stably into a range of different structures that serve a specific function such as, nanotubes of TMV coat protein that encapsulate the viral genome and aid in host infection[8], bacterial pili that help in gene transfer, and flagella that allow for bacterial locomotion[9], microtubules that are components of the cytoskeleton[10], nanostacks of 2-cys

peroxiredoxin that have holdase activity[11]. Such assembly of the protein units largely results from oligomerization of symmetric domains through complementary non-covalent interactions. Also, the highly complex and crowded environment *in vivo* provides the necessary entropic effects to guide specific assembly of proteins, as seen for proteins that self-assemble into filaments that are part of the cytoskeleton[12], protein fibres[13-19] and amyloid spherulites[14]. However, producing such stable architectures in the laboratory remains challenging as it is difficult to mimic the exact cellular environment and also, the experiments are carried out at dilute concentrations.

Tailoring protein building blocks, through a range of different techniques, have provided an alternate route to produce multiprotein structures *in vitro*[20]. Design strategies currently reported have relied mainly on modification of natural domains and/or incorporation of specific secondary structure elements through genetic modification or through chemical linking[7]. Protein tectons with an ability to self-assemble into nanometric structures have provided an alternative for the rational design of recombinant constructs that can self-assemble in response to a trigger[21]. Also, they can be expressed in large quantities and offer the ease of site-directed modifications and some of them offer the advantage of thermal stability, such as the LSMα from *Methanobacterium thermoautotrophicum*[22].

Peptides that self-assemble into extended nanostructures in response to physiological trigger have proven to be excellent candidates for production of functional hybrid materials[23]. In another approach, mixing of macromolecules with small molecules has led to production of macroscopic structures from molecular self-assembly through cooperative behavior. An important example is the production of large membrane structures by mixing the polysaccharide hyaluronic acid with peptide amphiphiles[24]. Xu *et al.* (2013) showed that

aggregates of amphiphilic peptides mixed with cytosol proteins bind to the protein in drastically different ways depending on the scale of peptide assemblies[25]. In more recent work, fibril-forming peptide motifs were successfully exploited as nano-scaffolds to which different proteins can be linked to create functional biomaterials, such as nanofibers with enzymatic activity and nanofibers with antigenicity for enhanced immune response [26]. Self-assembled peptide nanostructures as scaffolds present an attractive tool to template assembly of complex proteins due to ease of use, customization and control of assembly *in vitro*. This chapter discusses the strategy to generate well-ordered extended nanoscale architectures of *h*Prx3 protein, templated through self-assembling peptide sequences.

5.2 Peroxiredoxin a model self-assembling protein

Peroxiredoxins (Prxs), a family of enzymes that fix free peroxide in the cells and protect them from oxidative damage, exist as obligate homodimers that further self-assemble into toroidal oligomers[27]. Some Prxs can be triggered to form a range of different supramolecular assemblies[28], such as clusters[29], cages[30], interlocked rings[31], toroidal stacks[32-36] and short nanotubes[37]. The toroid form of the protein is a responsive assembly of the protein dimers triggered by reducing conditions[27]. The radial symmetry of the associated dimers gives rise to the toroidal face comprising of five or six repeating units of the protein dimer. The toroid-shaped oligomer is a generic structure of peroxiredoxin proteins with varying ring sizes and provides a large surface area of contact including the cavity of the toroidal pore. Amongst the different peroxiredoxins characterized, the dodecamer toroid-shaped quaternary conformation adopted by the mammalian Prx3 protein is the largest, with an external and internal diameter of 14 nm and 7.5 nm respectively[38,39,40] (Figure 5.1).

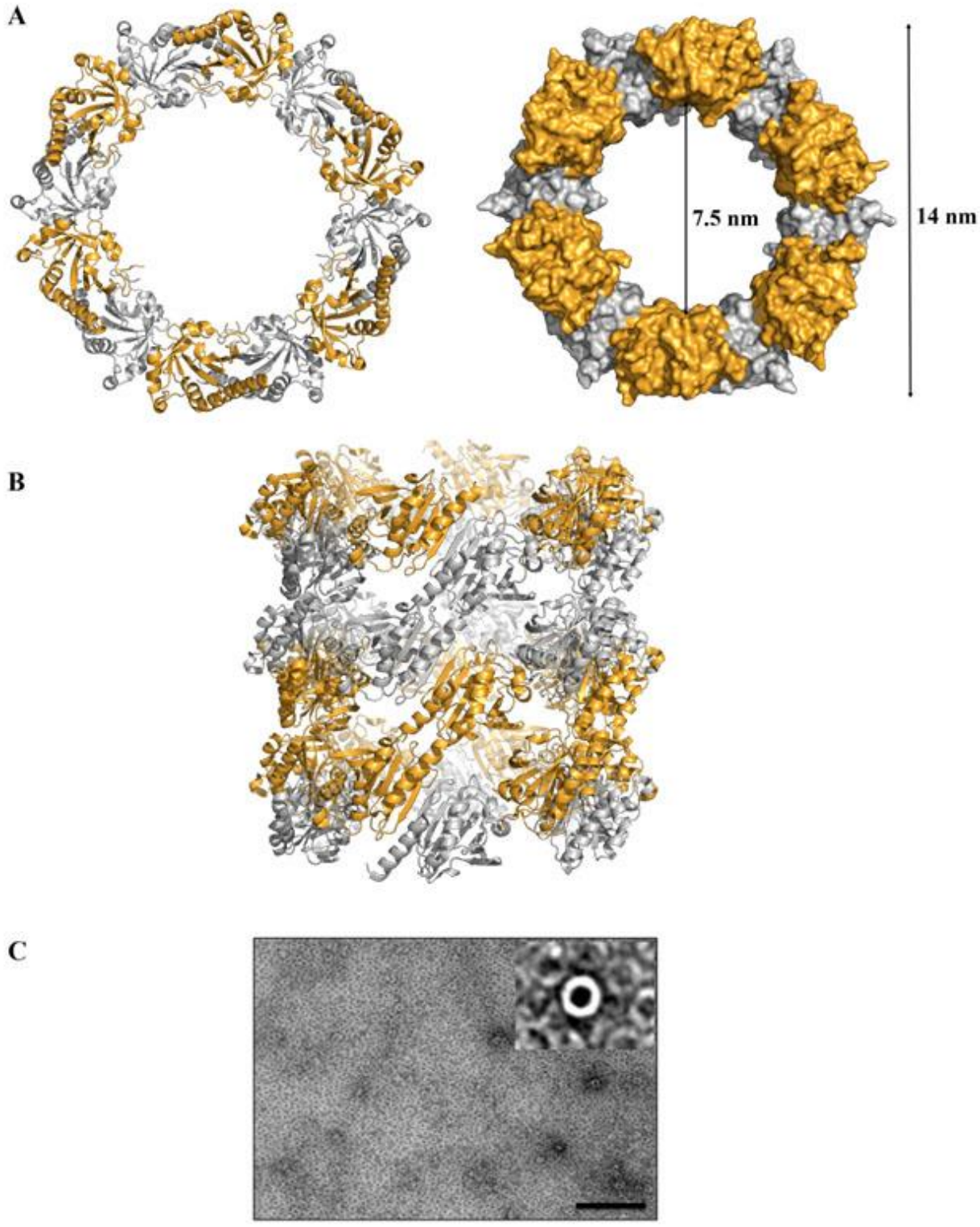


Figure 5.1: *Crystal structure and TEM of hPrx3: (A) Cartoon image of hPrx3 dodecamer (PDB 5JCG) showing the inner and outer diameter of the ring structure; (B) Cartoon image of hPrx3 toroidal stack observed in protein crystal at alkaline pH [41]; (C) TEM image shown was taken after dialyzing cleaved hPrx3 in pH 8.0 reducing buffer. (Scale bar is 100 nm).*

This feature of the Prx ring has been exploited for the production of toroidal nanostacks of peroxiredoxin from *S. mansoni* functionalized with gold nanoparticles[42]. An interesting feature observed in Prxs is the ubiquitous self-assembly of the toroids into short nanostacks of upto 500 nm in length under acidic conditions. A recent high-resolution cryo-EM structure of *h*Prx3 toroidal stacks showed that the stacking of the toroids is facilitated through a network of hydrogen bonds between the surface exposed α2, α6 and β2 motifs that lock the toroids into the high molecular weight stacked architecture[43,39]. In a more recent study, it was revealed that short toroidal stacks of Prx3 can form under alkaline pH under the highly crowded environment within the protein crystal. In the crystal structure the toroidal stacks can be further seen laterally associated to each other[41]. The extended toroidal stacked assembly of Prx protein is posited to be a distinct assembly of the protein with holdase activity[29,39].

These properties of *h*Prx3 make it attractive as a protein tecton to be explored for production of non-biological supramolecular assemblies that could eventually find application as nanoscaffolds in bionanotechnology. As mentioned earlier, the Prx3 toroid provides a large number of surface exposed positively and negatively charged amino acid residues that can act as sites for attachment of linkers such as lysine to link polymers towards the creation of hybrid material. The hollow core of the protein toroids and extended nanotubes can be used as encapsulate molecules with a specific activity. Furthermore, the toroid conformation is redox sensitive, thus provides a suitable control for targeted delivery and release of therapeutic molecules *in vitro* and their application as drug delivery vehicles. Modification of surface chemistry of the toroid through the introduction of specific residues like cysteine can allow for the production of non-biological architectures structures with redox sensitivity. Furthermore, incorporation of unnatural amino acids (UAA) is another prospective application that can be used for to produce protein architectures bearing antibiotic activity.

For example, the UAA phenylglycine a homolog of phenylalanine bearing aromatic properties and is found in different antibiotics such as Apalcillin and Mezlocillin.

5.3 Strategy for templating higher order assembly of *h*Prx3

Electrostatic surface potential analysis performed on the crystal structure of *h*Prx3 (PDB 5JCG) using pymol software (www.pymol.org) showed polar and charged residues richly distributed on the surface (Figure 5.2). The homo-dodecamer toroid of *h*Prx3 contains six repeating dimer units of the protein. The radial symmetry between the associated dimers generates a toroid shaped scaffold with four different surfaces. The four different facets of the toroid are easily accessible and provide a large surface area for the association of the ring with extended peptide nanostructures, and for the ring to ring association. Analysis of the protein sequence using the protein property calculator tool (Expasy and http://protcalc.sourceforge.net) revealed that under neutral to alkaline conditions the protein is highly charged with a theoretical net charge of -4 at pH 7 and a net charge of -8 at pH 8.

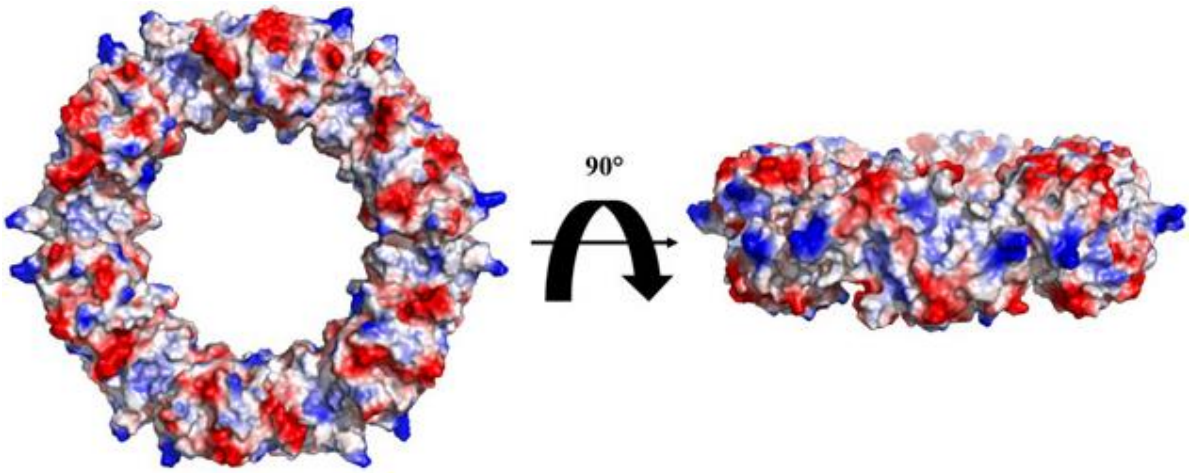


Figure 5.2: *An electrostatic surface potential model of hPrx3 (PDB 5JCG) generated at pH 8.0, showing the distribution of charged residue. Positively charged residue are shown in color blue and negatively charged residues are shown in the color red.*

As described in the previous chapter, the peptide sequences studied were comprised of one or more charged amino acid residues and the overall net charge on the peptide sequence is +1 or -1 at near neutral pH. Also, the side chain group of the charged lysine residue (K), the hydroxyl group of the serine residue (S) and the carboxamide group of the asparagine residue (N) within the β-continuous homodimer interface of Prx3 protein are seen protruding outwards from the β-sheet surface (Figure 4.1).

The self-assembled extended peptide fibrillar structures thus provide a large surface area of non-covalent contact points that is accessible to the *h*Prx3 toroids. It was therefore hypothesized that mixing the extended peptide nanostructures with *h*Prx3 protein toroids, would result in association of *h*Prx3 toroids with the peptide nanostructures in a specific manner to form extended peptide-protein assemblies.

5.4 Production of *h*Prx3 protein nanotubes

The *h*Prx3 protein was expressed and purified as described in chapter 4 (Section 4.10.1). The N-terminal 6*His-tag was cleaved from the fusion protein through site-specific cleavage by *rTEV* protease and further purified (Section 2.5.6.2 and 2.5.6.3). The purity of the product was judged by SDS-PAGE analysis and complete removal of 6*His-tag was confirmed by liquid chromatography coupled with mass spectrometry (Appendix 1, Figure A1). The mixing experiments were carried out in HEPES, NaCl buffer at pH 8.0 buffer containing TCEP (Section 2.2.6) as the protein is stable in the buffer and adopts the toroid shaped quaternary conformation. Also, the self-assembly of the peptide was found to unaffected in this condition (Figure 5.3 A). Unless otherwise stated, the peptide solution was prepared by dissolving the peptide acetate powder in pH 8.0 buffer above the critical assembly concentration, vortexed to mix and equilibrated for 24 hours at 4°C prior to mixing.

5.4.1 Mixing Prx3 peptide with *h*Prx3 protein

The tagged *h*Prx3 protein sample stored at -80°C was thawed at room temperature and dialyzed into pH 8.0 buffer containing TCEP overnight (Section 2.5.7.2) and concentrated to 10 mg/mL (Section 2.5.5.3). The peptide acetate was mixed with 6*His *h*Prx3 protein, as described in section 2.5.9. TEM analysis of the peptide-protein mix revealed extended nanotubular assemblies with *h*Prx3 toroids scaffolded onto peptide fibrils (Figure 5.3 B, C). Aside, from the extended nanotubular structures, some free Prx3 peptide fibrils and *h*Prx3 toroids were also observed on the EM grid which suggested a heterogeneous mix of nanotubes, fibrils, and free toroids.

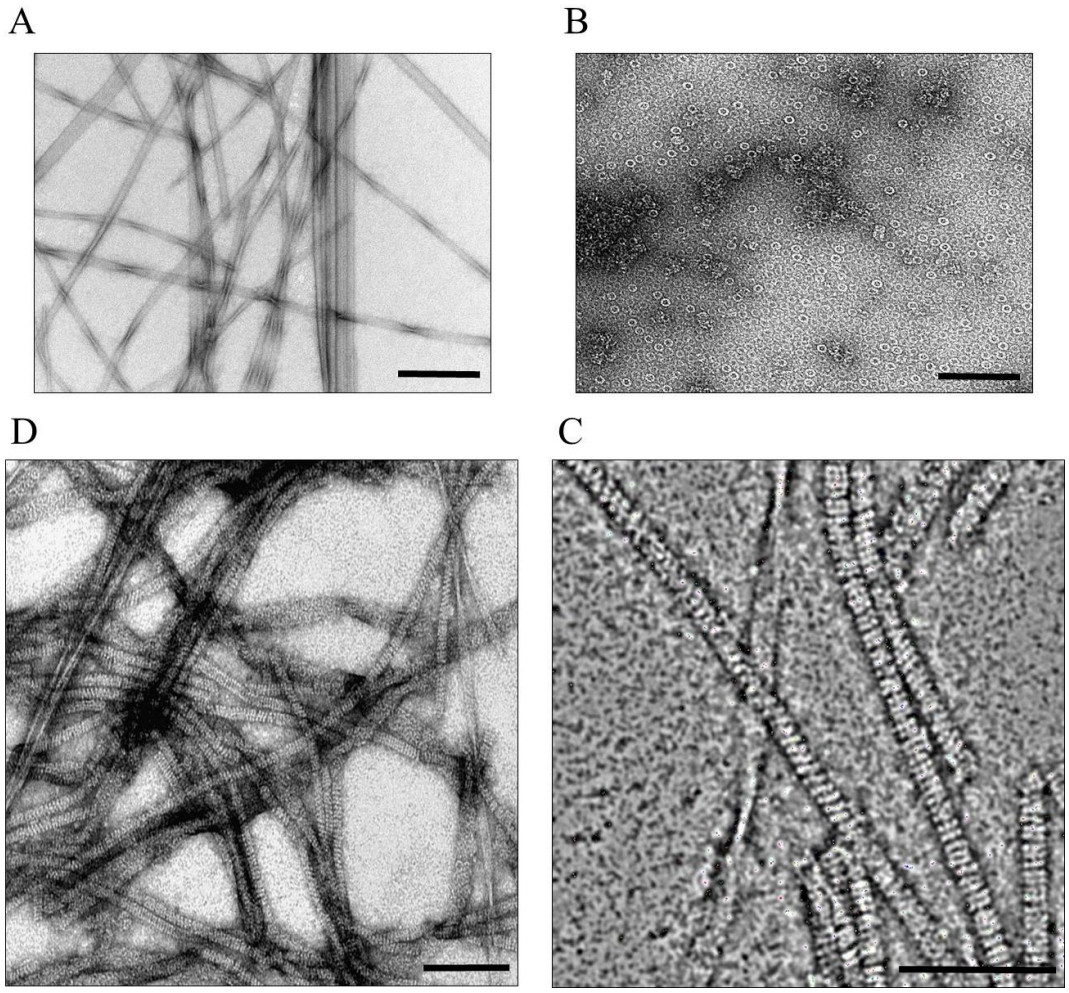


Figure 5.3*: TEM showing the extended nanotubular assembly of 6*His hPrx3 templated by Prx3 peptide nanotapes.*

(A) Control sample of peptide acetate in buffer at 5 mg/mL.

*(B) Control sample of 6*His hPrx3 protein in buffer at 0.1 mg/mL.*

*(C) Extended nanotubes formed upon mixing Prx3 peptide tapes to 6*His hPrx3 protein toroids (0.1 mg/mL) equilibrated for 24 hours at 4°C.*

(D) Zoomed EM micrograph produced using ImageJ by applying bandfass filter, shows the details of the protein toroids arranged into extended nanostacks and free peptide nanotape.

Note: Experiments were conducted in 20 mM Hepes, 150 mM NaCl at pH 8.0. (Scale bar is 100 nm)

To examine the significance of the peptide tag in the nanotubular assembly of the *h*Prx3 protein, the experiments were repeated with the cleaved *h*Prx3. The presence of extended tubular structures of nano- and microscale lengths similar to those observed with tagged protein demonstrated that the specific assembly of the protein toroids with the peptide tapes was not 6*His-tag dependent (Figure 5.4).

TEM further revealed that the tubular nanostructures are extended with many of them spanning across the imaged cross-section of the EM grid. Measurements performed on the TEM micrographs using ImageJ software revealed that the nanotubular structures were heterogeneous in length, with an average overall length between 1 to 2 micrometers. The width of the extended nanotubes was measured as a distribution between 10 to 15 nm[44].

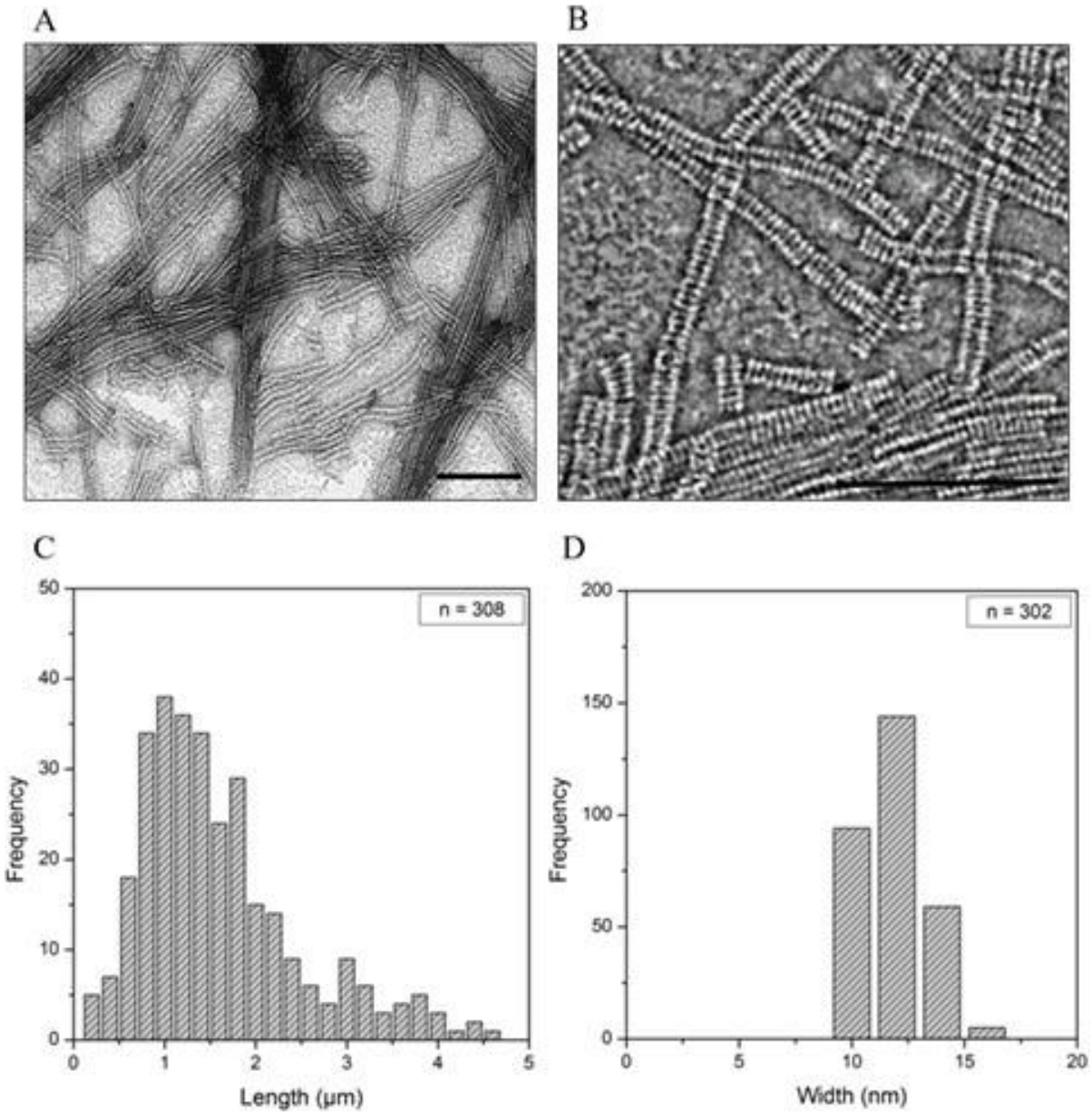


Figure 5.4*: TEM showing the extended nanotubes of hPrx3 templated by Prx3 peptide nanotapes: (A) Extended nanotubular structures formed on mixing Prx3 peptide mixed with hPrx3 protein; (B) Zoomed EM micrograph produced using ImageJ; (C) Histogram showing the distribution of nanotube length and nanotube width (D) collected from 308 and 302 particles respectively. (Scale bar is 100 nm)*

Note: The concentration of peptide is 5.0 mg/mL and protein concentration is 0.1 mg/mL (0.0003 mM) with a molar ratio is 15000 peptide molecules:1 protein toroid. Experiments were conducted in 20 mM Hepes, 150 mM NaCl, 3 mM TCEP at pH 8.0.

To assess whether the assembly of *h*Prx3 protein with Prx3 peptide was specific, two types of strategies were undertaken:

(i) *Mixing Prx3 peptide with other toroid shaped proteins.*

(ii) *Mixing other self-assembling peptide sequences with hPrx3 protein.*

5.4.2 Mixing Prx3 peptide with *Ap*Tpx and Lsmα

To examine the interaction of the peptide tapes with other toroid shaped proteins, the experiment was repeated with a decamer peroxiredoxin derived from the thermophile *Aeropyrum pernix* (*Ap*Tpx)[45] and the unrelated toroid shaped heptamer protein Lsmα from *Methanobacterium thermoautotropicum*[22].

Mixing experiments were performed in solution conditions most conducive for the protein: In Hepes pH 8.0 buffer for *Ap*Tpx and in PBS pH 8.0 for Lsmα (Section 2.2.6). The overall theoretical charge on protein in the experimental conditions is -4.5 (*Ap*Tpx) and -2.5 (Lsmα). No specific extended nanotubular assembly was observed on mixing the two unrelated proteins with the peptide tapes (Figure 5.5).

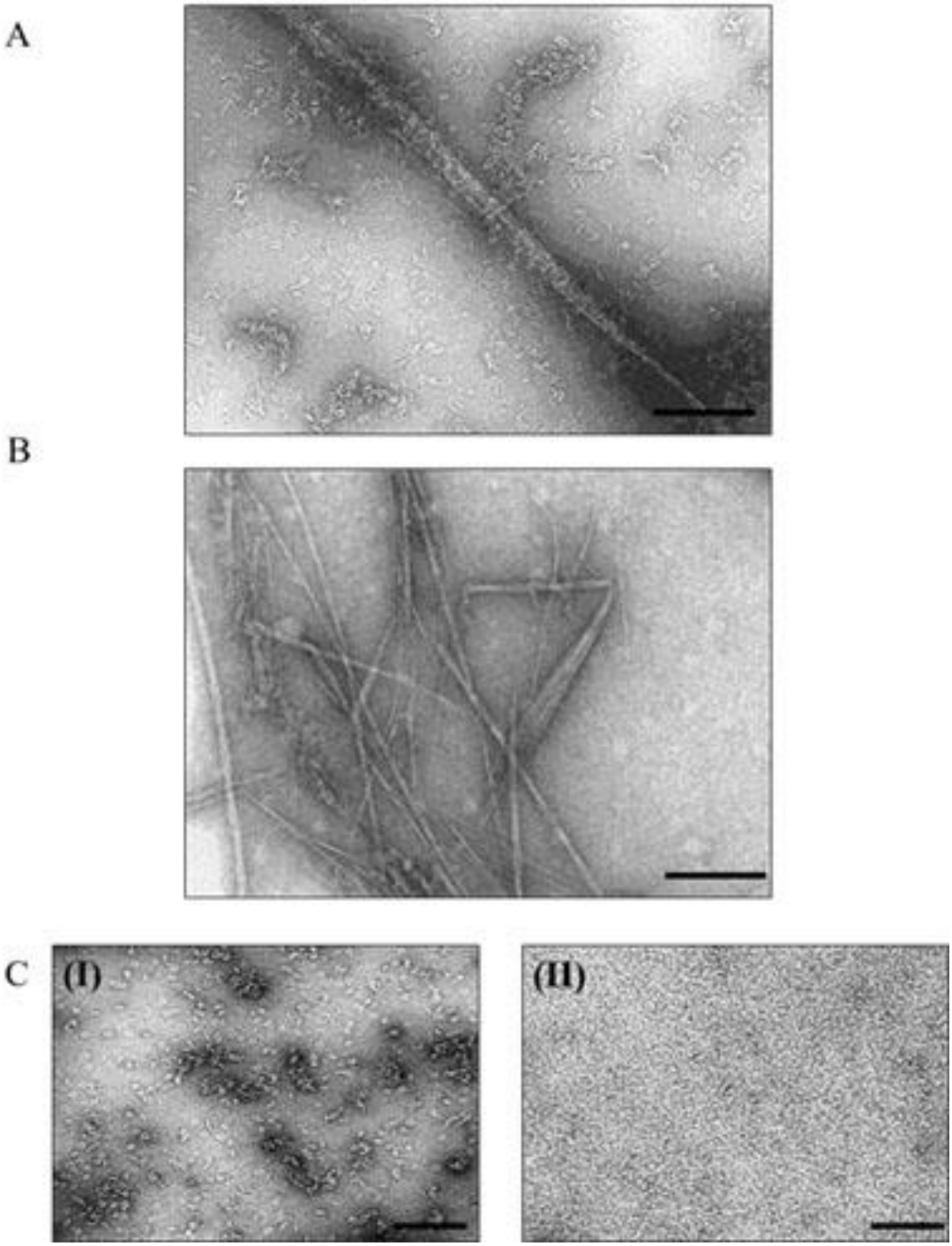


Figure 5.5: *TEM of Prx3 peptide nanotapes mixed with other toroidal proteins: EM micrograph of Prx3 peptide mixed with ApTpx decamer toroids (A) and peptide mixed with Methanobacterium thermoautotrophicum LSMα heptamer toroids (B). Control sample show mostly individual decamer toroids for both ApTpx (CI) and Methanobacterium thermoautotrophicum Lsmα (CII) at 0.1 mg/mL. The peptide acetate concentration in the sample mix is 5 mg/mL. (Scale bar is 100 nm)*

The solution conditions for the experiment were as follows:

(A) Hepes buffer: 20 mM Hepes, 150 mM NaCl, 3 mM TCEP pH 8.0

(B) PBS buffer: 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, 2.7 mM KCl, 140mM NaCl pH 8.0

The roman numerals show the control protein samples: (I) ApTpx in Hepes pH 8.0 buffer and (II) Lsmα in PBS pH 8.0 buffer.

Control samples of protein in their respective buffers revealed mostly individual protein toroids in both the cases. Control peptide samples showed extended nanotapes like assemblies as observed in control pH 8.0 buffer. The results thus suggested that nanotubular assembly observed was specific to the Prx3 protein.

5.4.3 Mixing different peptide tectons with *h*Prx3

To investigate the significance of altering the peptide sequences on the peptide-protein assembly, the interaction of *h*Prx3 protein with self-assembling peptide sequences other than Prx3 peptide was studied. Peptide sequences described in chapter 3 (Section 3.3) were prepared in the experimental buffer above the critical concentration of assembly and mixed with the *h*Prx3 protein. The conditions of assembly observed with the Prx3 peptide were maintained for the series of peptide sequences investigated for the specificity of the interaction with the protein.

TEM of the protein-peptide mixes revealed that all the peptide sequences trialed induced hierarchical micron-long tubular stacks of the *h*Prx3 protein, with dense bundles of nanotubes observed with β-Lac peptide mixed with protein, individual nanotubes along with some bundles of nanotubes with DAPDC peptide and mostly individual nanotubes with UmuD' peptide (Figure 5.6). The results suggested the assembly is specific to *h*Prx3.

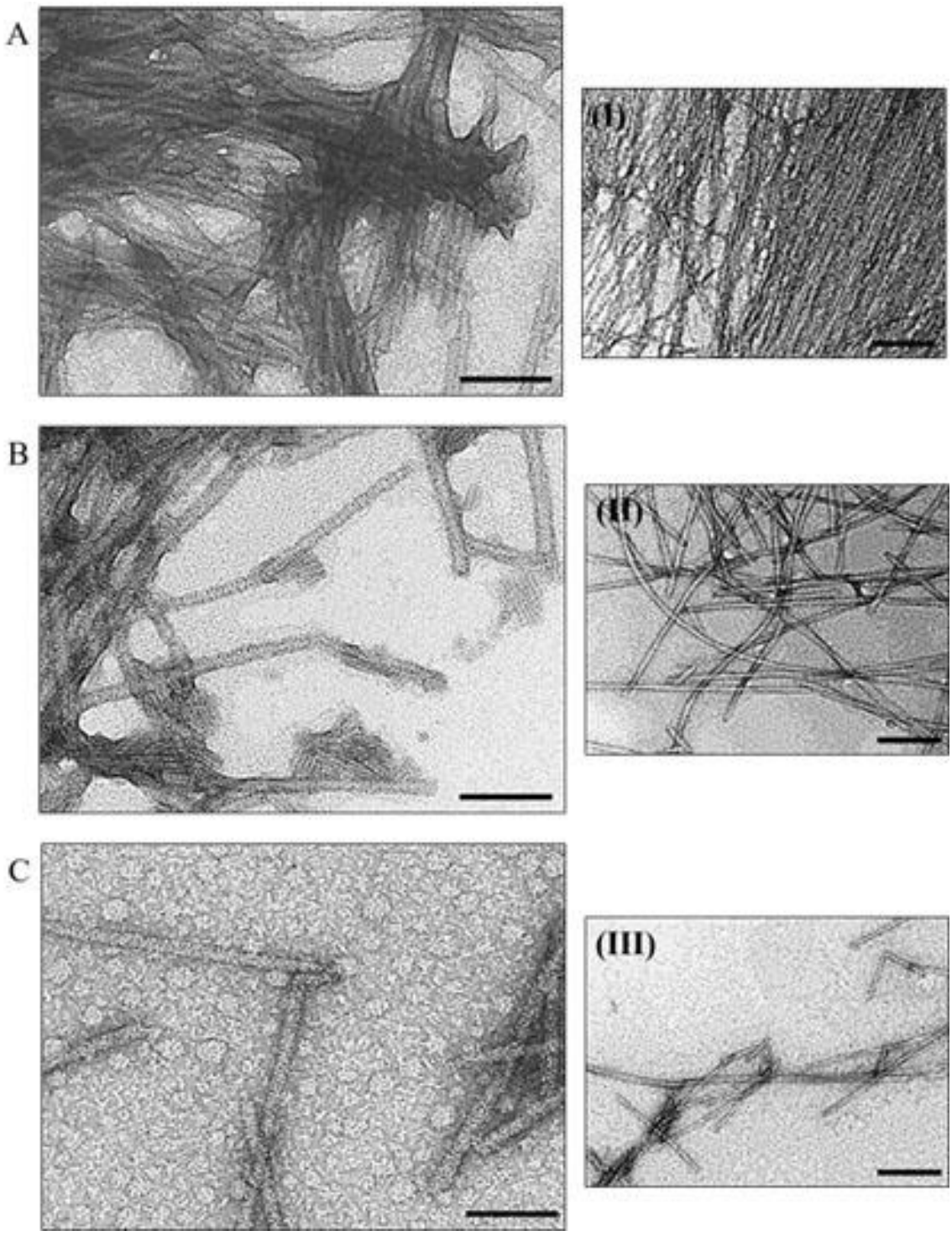


Figure 5.6: *TEM showing the nanotubular assembly of hPrx3 protein induced by different peptide sequences: (A) hPrx3 + β-Lac peptide; (B) hPrx3 + DAPDC peptide; (C) hPrx3 + UmuD' peptide. Control peptide acetate samples in reducing pH 8.0 buffer show self-assembled nanostructures: (I) β-Lac peptide (10 mg/mL); (II) DAPDC peptide (70 mg/mL); (III) UmuD' peptide (100 mg/mL). Protein concentration is 0.1 mg/mL. (Scale bar is 100 nm)*

5.5 Production of decorated peptide tapes and protein nanosheets

Variants of the Prx3 peptide as described in Chapter 4 mixed with the *h*Prx3 protein also induced extended tubular assembly of the protein toroids similar to that described for parent Prx3 peptide. In the case of H3Y variant containing the specific sequence change with the histidine residue replaced by tyrosine residue, mixed with protein toroids resulted in a non-specific assembly, with peptide tapes decorated with protein toroids (Figure 5.7 B).

As described in chapter 4 (Section 4.7), the H3Y peptide variant self-assembles into a heterogeneous mix of thin and wide extended peptide tapes. The polymorphic peptide tapes with varying widths present different type of scaffolds to which protein toroids can associate. In the case of thin peptide tapes, relatively smaller surface area is available to the protein toroids thus, the protein toroids associate laterally which result in nanotube assembly. The wider peptide tapes present a large surface area of contact that is available to the protein toroids. This explains the non-specific assembly of the toroids decorated onto the extended peptide tapes. Therefore, to examine the non-specific assembly of the toroids with the variant peptide, mixing studies were carried out at large molar excess of the peptide that favored wide tape like assembly.

TEM of variant peptide-protein mix containing 150 fold higher concentration (w/v) of the peptide to protein revealed nanotubular structures, with protein toroids, found laterally associated to the extended peptide tapes (Figure 5.7 A). For the sample containing a 250 fold excess of the peptide mixed with the protein, TEM revealed peptide nanotapes decorated with protein rings (Figure 5.7 B). In the case of a peptide-protein mix prepared to contain 500 fold excess (w/v) of the peptide mixed with protein, the *h*Prx3 toroids were found to laminate on top of the wide extended peptide tapes forming protein nano-sheets (Figure 5.7 C).

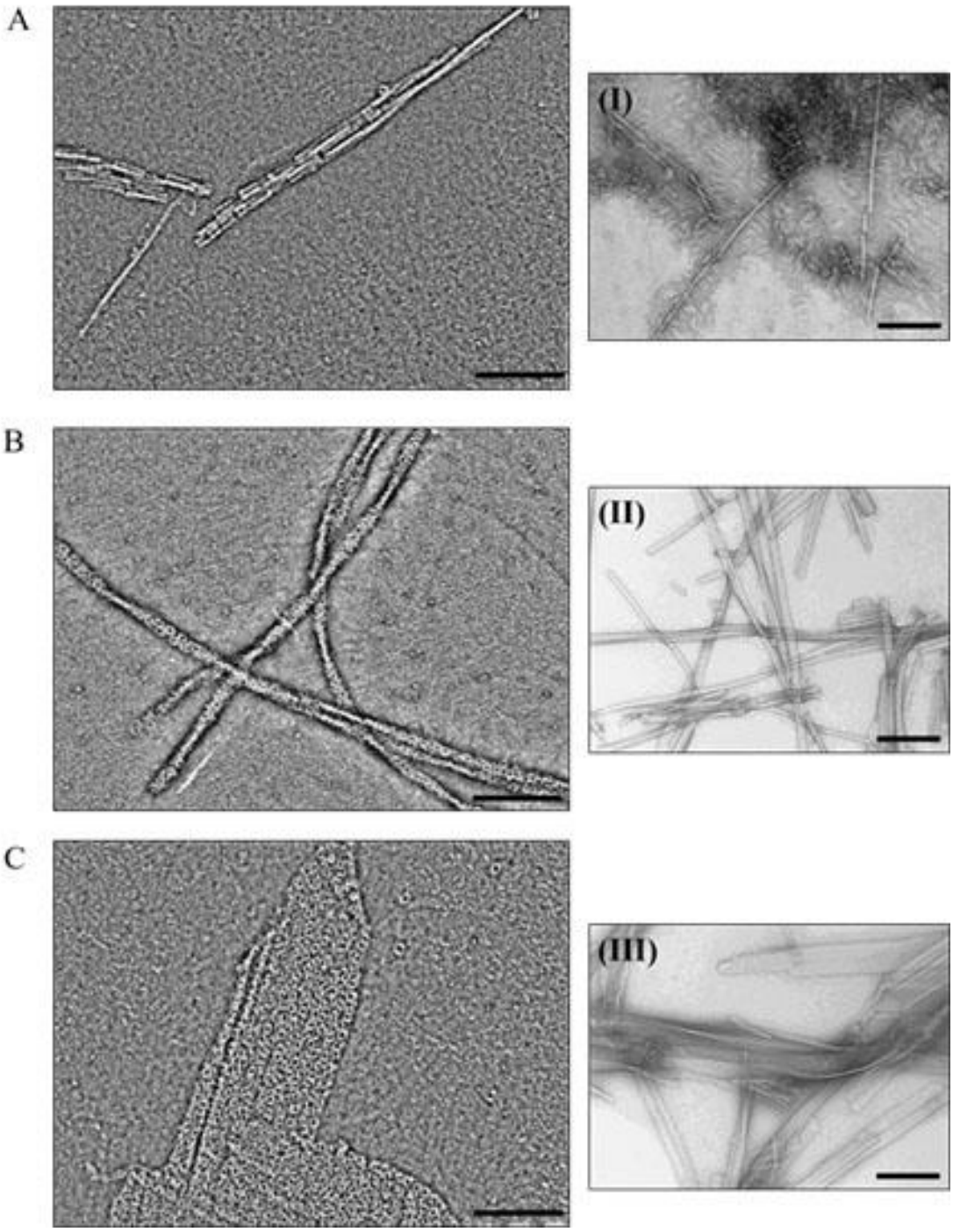


Figure 5.7: *TEM of variant peptide acetate H3Y mixed with hPrx3 at different concentrations: (A) H3Y peptide mixed at 150 fold higher excess to protein show nanotubes and peptide fibrils; (B) H3Y peptide mixed at 250 fold higher excess to protein show decorated peptide tapes; (C) H3Y peptide at 500 fold higher excess to protein show protein nanosheets. (Scale bar is 100 nm).*

Note: Control peptide acetate sample prepared in reducing pH 8.0 buffer show peptide nanostructures: (I) Nanofibrils at 15 mg/mL; (II) Nanotapes at 25 mg/mL; (III) Wide nanotapes at 50 mg/mL. (Scale bar is 200 nm)

5.6 Factors affecting peptide-protein assembly

In order to probe the factors that control the different modes of interaction of *h*Prx3 toroids with peptide tapes, different solution conditions, and concentrations of peptide were tested. Previous studies on *h*Prx3 protein showed that the ionic strength of the buffer influences the self-assembly of the protein toroids into nanostacks[36]. Furthermore, pH is known to play an important role in the stacking of the protein toroids into extended tubes with acidic pH favoring the assembly of multiple toroids[35,42].

5.6.1 Peptide-protein assembly at different pH conditions

Solution conditions such as pH can have a substantial effect on the quaternary organization of proteins and peptides. In the case of *h*Prx3, the protein toroids are stable at acidic and alkaline pH conditions. Furthermore, acidic pH < 6.0 favors the assembly of multiple toroids into short nanotubes [36]. Therefore, to study the effect of pH on the assembly of peptide tapes with protein toroids, the peptide-protein mixes were prepared in different pH buffers. As, the theoretical pI of *h*Prx3 is 6.8, a pH range of 4.0 to 9.0 was chosen.

The self-assembly of *h*Prx3 toroids and the Prx3 peptide was found to be stable across the 4.0 to 8.0 pH range (Figure 5.8 A, B). This was apparent from the extended nanotubular assembly observed for the peptide-protein mixes prepared in pH 8.0 buffer and dialyzed into pH 6.0 buffer and pH 4.0 buffer overnight. In the case of a peptide-protein mix dialyzed into pH 9.0 buffer, mostly short peptide nanotapes and aggregated protein were observed (Figure 5.8 C).

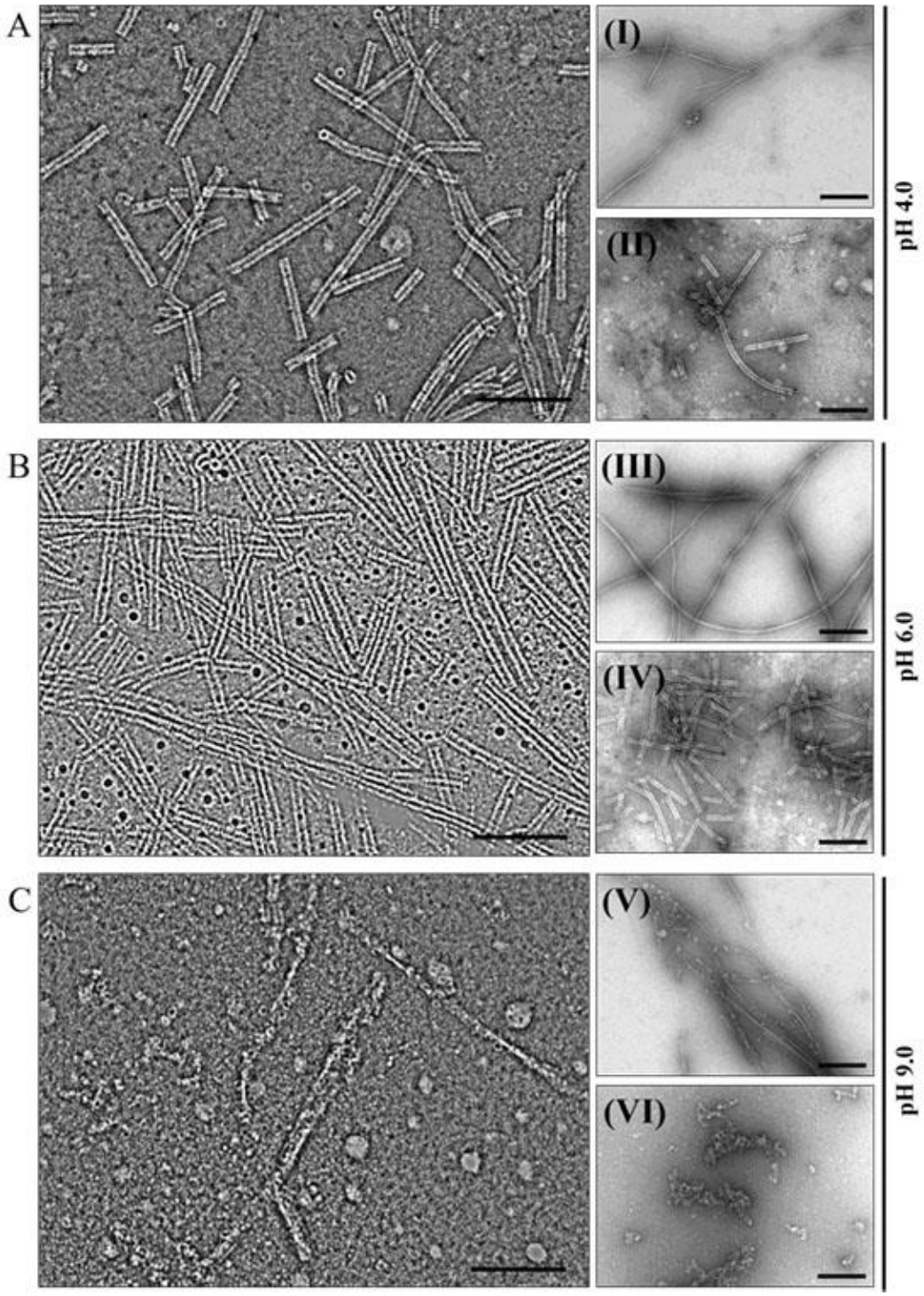


Figure 5.8: *TEM of the peptide-protein mix at different pH: Peptide-protein mix dialyzed into pH 4.0 buffer (A) and pH 6.0 buffer (B) that show extended nanotubes. (C) Peptide-protein mix in pH 9.0 buffer shows fewer peptide tapes. Peptide concentration in the mix is 5 mg/mL and protein concentration is 0.1 mg/mL. (For A,B and C, the scale bar is 100 nm and for I-VI the scale bar is 200 nm)*

The solution conditions for the experiment were as follows:

(A) pH 4.0: 20 mM sodium citrate, 150 mM NaCl, 3 mM TCEP

(B) pH 6.0: 20 mM MES, 150 mM NaCl, 3 mM TCEP pH 6.0,

(C) pH 9.0: 20 mM sodium borate, 150 mM NaCl, 3 mM TCEP pH 9.0

Control buffer: 20mM Hepes, 150 mM NaCl, 3 mM TCEP pH 8.0

The roman numerals show the control peptide acetates dissolved in pH 4.0 buffer (I); peptide acetate in pH 6.0 buffer (II); peptide acetate in pH 9.0 buffer. The control protein samples are at pH 4.0 (II), pH 6.0 (IV) and pH 9.0 (VI).

Measurements performed on imaged nanotubes for the peptide-protein mix under acidic pH were found to be larger in length than those formed by protein only in this condition. Control experiments performed with Prx3 peptide acetate in pH 4.0 buffer and pH 6.0 buffer showed mostly self-assembled peptide nanotapes consistent with those observed at pH 8.0. At pH 9.0, the peptide nanostructures were relatively shorter than those observed at pH 8.0 and acidic pH. These results also showed that the self-assembly of the Prx3 peptide was stable at pH 4.0. For protein only samples, the extended nanotubes of *h*Prx3 protein observed at acidic pH $\geq$ 6.0 was consistent with those reported earlier[36].

TEM of *h*Prx3 protein dialyzed into pH 9.0 buffer did not show any toroids, consistent with the SEC-SLS data showing the protein existed as a dimer (Figure 5.9). This result explained the absence of the nanotubular assembly for the protein-peptide mix at pH 9.0.

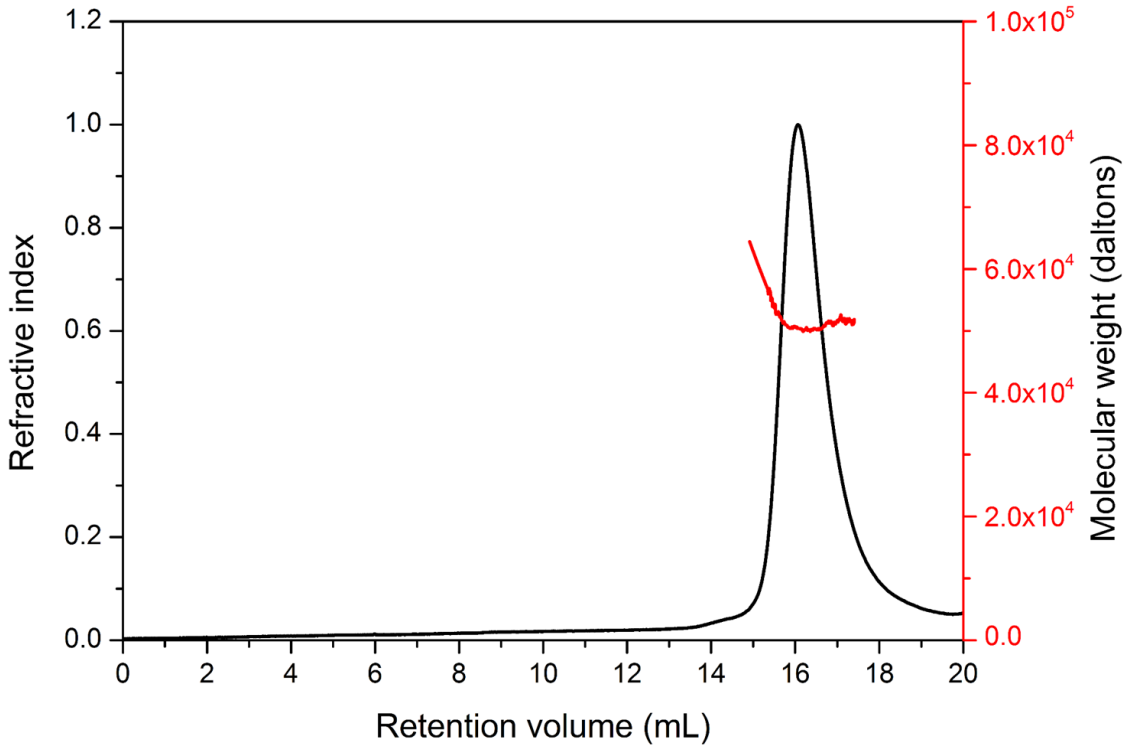


Figure 5.9: *The SEC-SLS trace of hPrx3 protein at 0.1mg/mL in reducing buffer pH 9.0, shows a single peak corresponding to the protein dimer.*

5.6.2 Peptide-protein assembly at high salt concentration

Samples of peptide-protein mix dialyzed into a buffer with varying amounts of sodium chloride concentration suggested that the peptide-protein assembly is disrupted by the high concentration of salt in the buffer. TEM of peptide-protein mix dialyzed into the high salt buffer (reducing pH 8.0 buffer + 500 mM sodium chloride) overnight showed mostly free peptide fibrils, some toroids associated non-specifically to peptide fibrils, a few nanostacks and free protein toroids (Figure 5.10).

The protein nanosheets produced using the H3Y peptide variant were also shown to disassemble in buffer with high salt concentration (Appendix 1, Figure A2). These results suggested that the assembly of *h*Prx3 toroids with Prx3 peptide and peptide variant H3Y is largely of electrostatic nature.

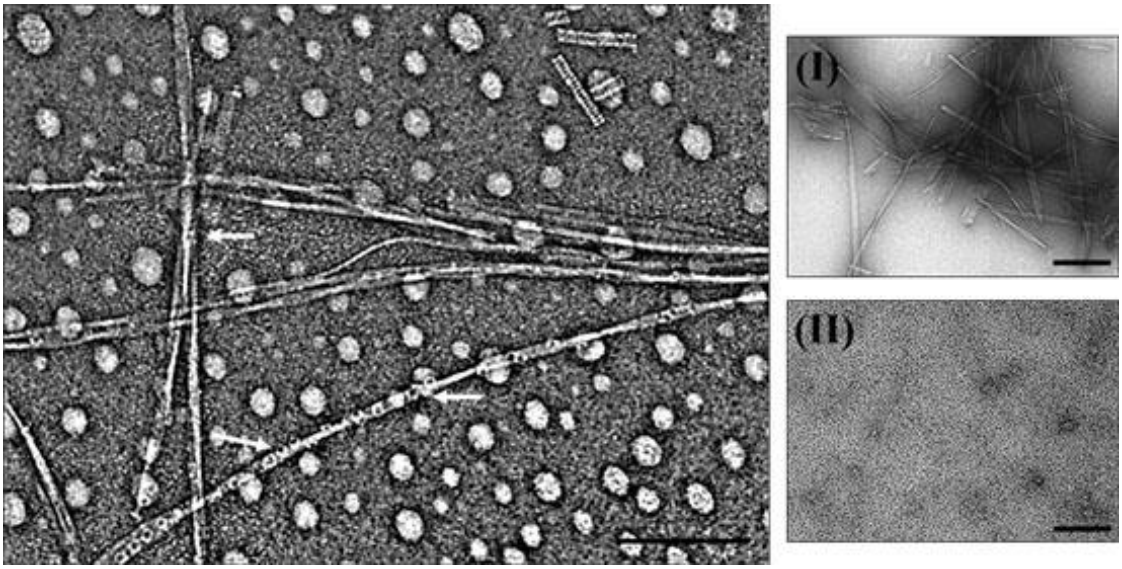


Figure 5.10: *TEM of the peptide-protein mix at high salt concentration: Peptide-protein in reducing pH 8.0 buffer with 150 mM NaCl dialyzed into high salt pH 8.0 buffer overnight, showing mostly free peptide fibrils, short nanotubes, and free protein toroids. The arrows point to the protein toroids associated with peptide-fibrils in a non-specific manner. (I) Peptide control sample at 5 mg/mL shows extended nanostructures and (II) Control protein sample concentration at 0.1 mg/mL shows mostly protein toroids. (Scale bar is 100 nm)*

5.6.3 Peptide-protein assembly at varying ratio of peptide to protein

Peptide-protein mixes prepared at dilute concentration of the Prx3 peptide and H3Y peptide variant at 1 mg/mL in reducing pH 8.0 buffer equilibrated for 24 hours revealed mostly free toroids and some toroids laterally associated to the peptide fibril. However, negatively stained EM micrographs of peptide and protein mixes equilibrated over 48 hours revealed short toroidal stacks and protein toroids laterally associated to peptide tapes (Figure 5.11). As previously described in chapter 4, the peptide sequences under dilute conditions exhibit slow kinetics of self-assembly. Also, the extent of self-assembly of the peptide is much lesser, which explains the result for dilute peptide-protein mixes.

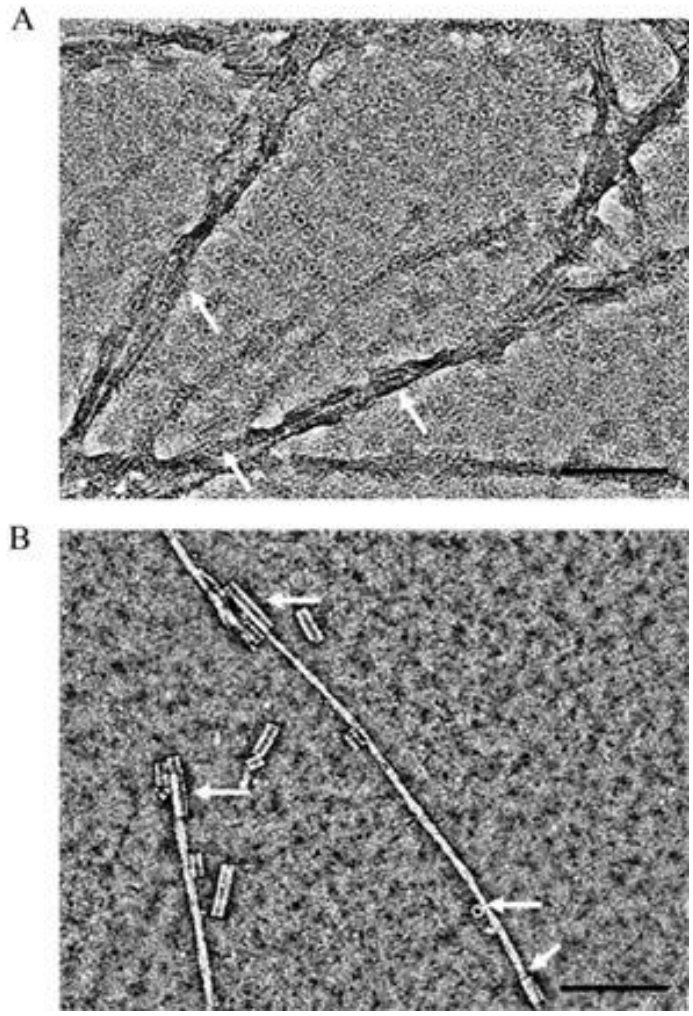


Figure 5.11: *TEM micrographs of hPrx3 protein toroids mixed with Prx3 peptide (A) and H3Y variant (B) in HEPES, NaCl buffer pH 8.0 equilibrated for 48 hours. Arrows show hPrx3 protein toroids laterally associated to extended peptide tapes. (Scale bar is 100 nm).*

The results highlighted that the formation of extended peptide nanostructures is necessary to template nanotubular assembly of *h*Prx3 protein toroids. To investigate the effect of the amount of peptide mixed with *h*Prx3 protein on the overall length of the nanotubular structures, the experiments were carried out at different concentrations of the peptide, keeping the protein concentration constant. Negatively stained TEM micrographs of the different Prx3 peptide and *h*Prx3 protein mixes showed an overall decrease in the size distribution of the nanotubes with a decrease in molar excess of peptide to protein (Figure 5.12).

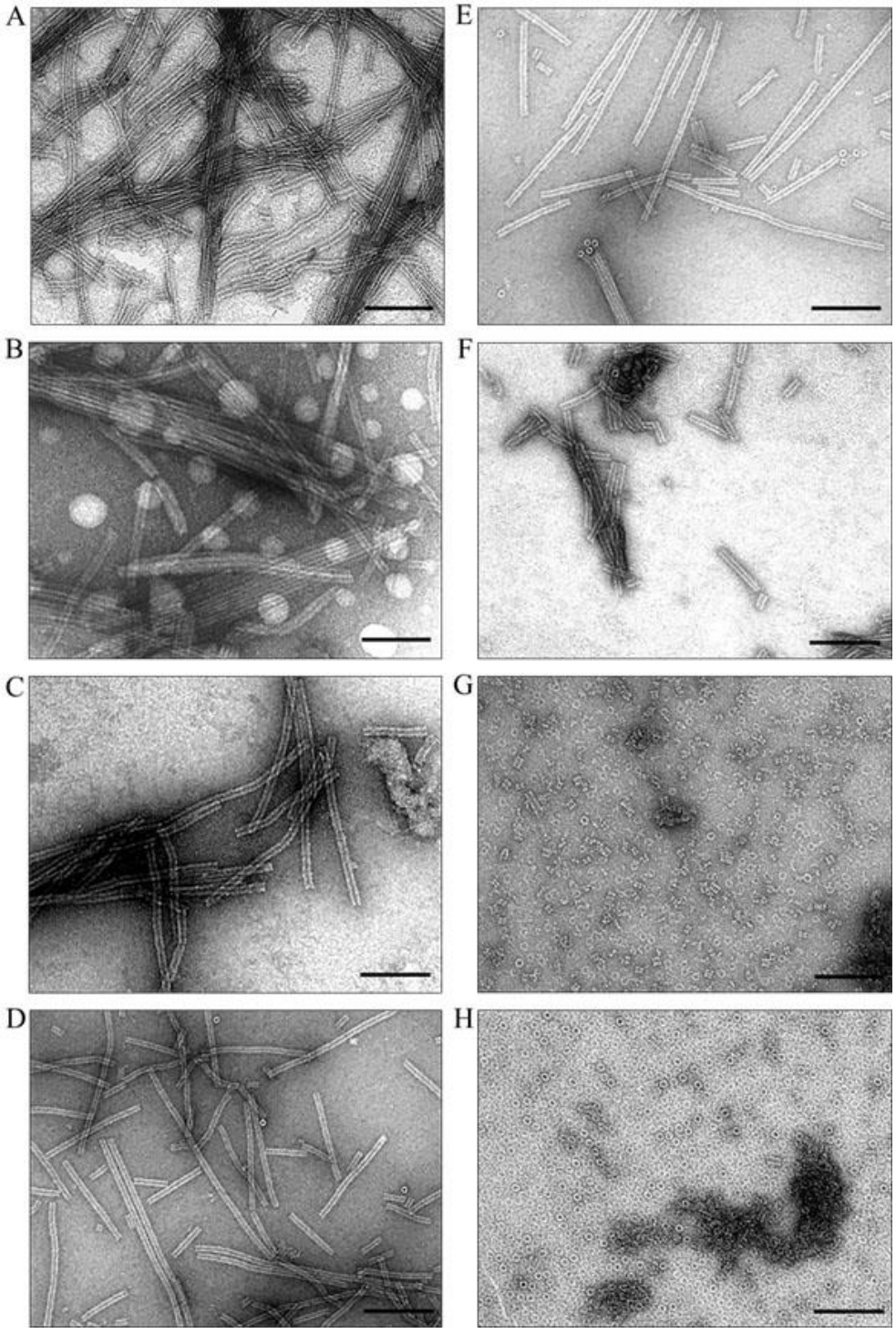


Figure 5.12: *Representative TEM micrographs of Prx3 peptide mixed with hPrx3 protein at (0.1 mg/ml) in reducing pH 8.0 buffer at different ratios. The final peptide concentration in the peptide-protein mix is: (A) 5mg/mL, (B) 2.5 mg/mL, (C) 1 mg/mL, (D) 0.5 mg/mL, (E)0.17 mg/mL, (F) 17 µg/mL, (G) 0.34 µg/mL (equimolar concentration to protein) and (H) 6.8 ng/mL (below equimolar concentration to protein). (Scale bar is 100 nm)*

Measurement of the nanotubular structures from different EM grids of the peptide-protein mixes performed using the ImageJ software, presented an overall decrease in nanotube length with decrease in the amount of peptide mixed with the protein (Figure 5.13). Thus, supporting the inference that varying the amount of peptide mixed with protein allowed some control on the overall length distribution of the nanotubular structures formed. The extended micron long nanotubular assembly of the *h*Prx3 protein described here have not been previously reported.

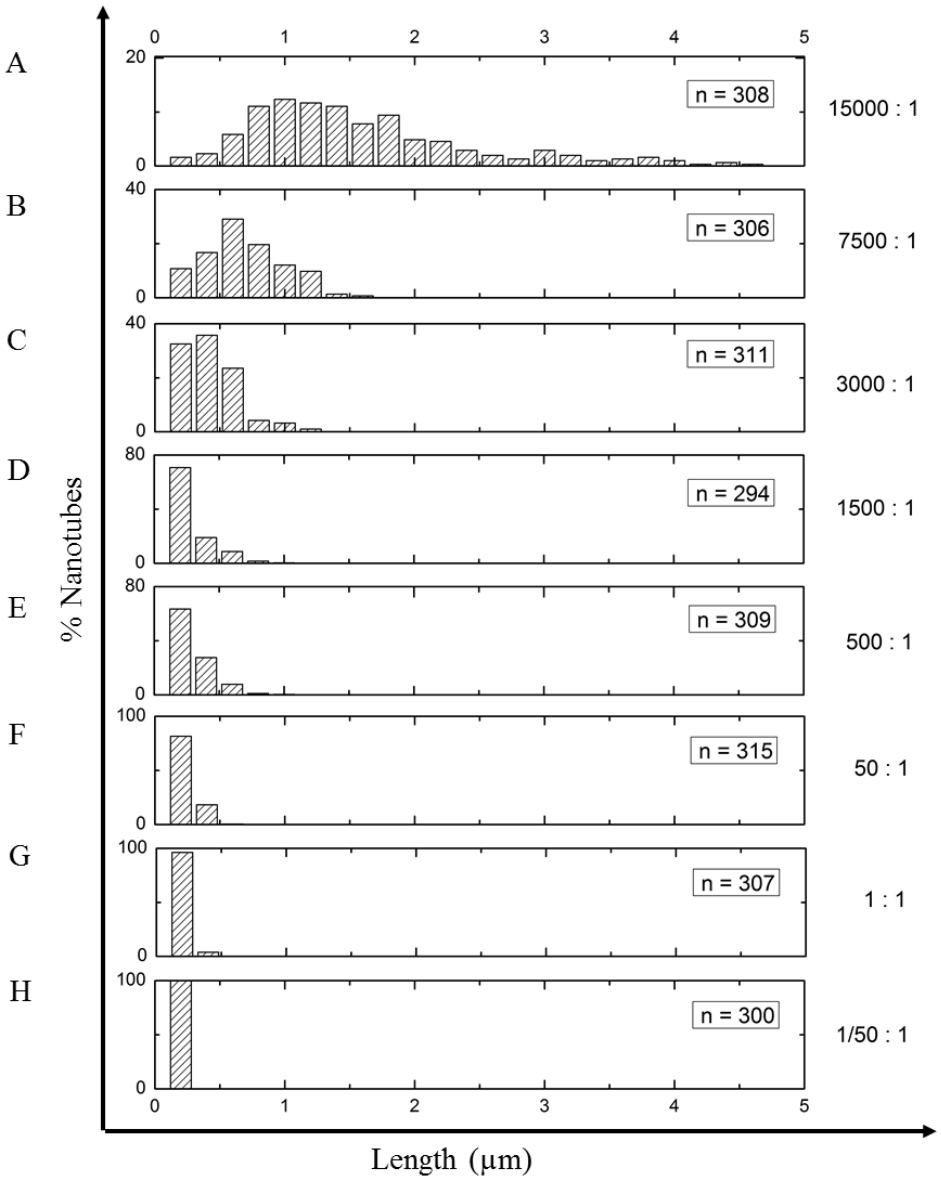


Figure 5.13: *Length distribution of nanotubes collected from ≥ 300 particle measurements performed using the ImageJ software. The ratios indicate the number of peptide molecules to protein for each of the peptide-protein mixes shown in Figure 5.12.*

The maximum length of nanotubes formed on mixing Prx3 peptide with hPrx3 toroids shows an overall decrease in the length of nanotubes relative to the amount of peptide concentration in solution.

The peptide-protein mixes prepared with the dilute concentration of the peptide mixed with the protein showed mostly short nanostacks consistent with the presence of free toroids (Figure 5.12 and Figure 5.13 panel F, G & H). These results indicated that at high concentrations of the peptide in solution (where extended peptide nanostructures are present) mixed with the protein produces extended nanotubular assembly. With the decrease in the amount of peptide mixed with protein in solution, the resultant nanotubular structures show a relative decrease in the overall size. Control samples of the peptide revealed mostly extended peptide nanotapes for the peptide acetate in buffer sample in the 5 mg/mL to 1 mg/mL range, though at a dilute concentration of 0.5 mg/mL mostly short peptide tapes were present. Below 0.5 mg/mL very few or no extended structures were observed on EM grid (Appendix 1, Figure A3).

5.7 Solution state characterization of peptide-protein assembly

To study the extended nanotubular assembly observed by TEM in solution, the peptide-protein mixes were studied as a function of the average hydrodynamic size by DLS[46,47]. DLS provides a fast and straightforward method for estimation of the mean size of nanostructures in solution[48].

Hydrodynamic size distribution by intensity was chosen as the method of analysis as the emphasis is on the species with the largest scattering intensity contributing to the overall result. DLS showed a decrease in the overall size of nanostructures in solution with the decrease of Prx3 peptide mixed with *h*Prx3 protein (Figure 5.14 A).

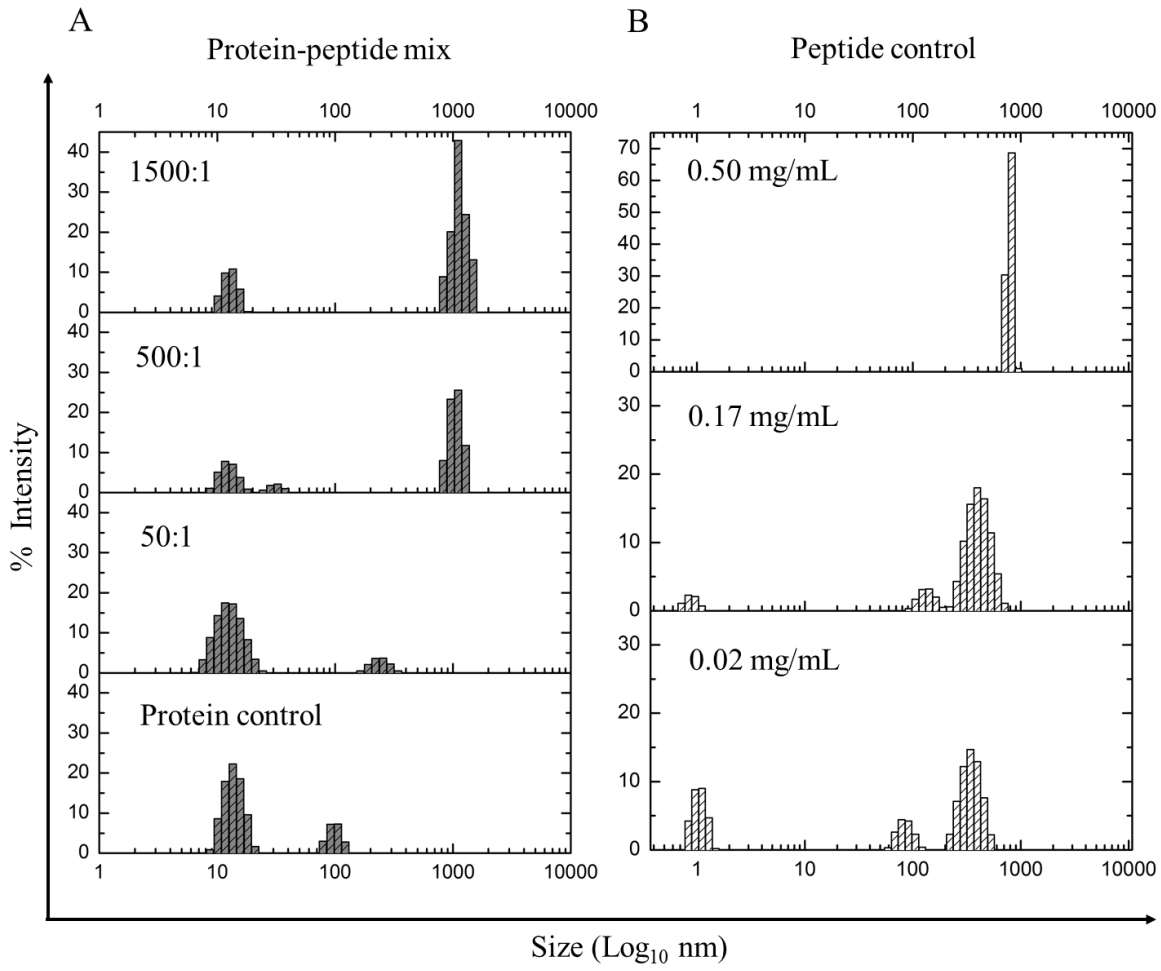


***Figure 5.14**: DLS of protein-peptide mixes in reducing pH 8.0 buffer plotted as size distribution by intensity. (A) Prx3 peptide and hPrx3 protein mixes with different concentrations of peptide mixed with protein. The final concentration of the peptide in the mix is: 0.50 mg/mL (1500:1); 0.17 mg/mL (500:1); 0.02 mg/mL (50:1) and protein control is hPrx3 at 0.1 mg/mL. (B) Peptide control samples: peptide acetate in reducing pH 8.0 buffer. The peptide-protein mix and control samples were incubated at 4°C for 24 hours prior to analysis.*

The intensity by size data for the peptide-protein mix containing higher fold excess of peptide mixed with protein (1500:1) revealed a polydisperse sample consistent with the presence of large species with an average size of ~1000 nm. DLS of the peptide-protein mix with lower fold excess of the peptide to protein (500:1), showed a decrease in the intensity by size which

suggested the decrease in the number of large nanostructures in solution. For the peptide-protein mix with a tenfold lower excess of the peptide mixed with protein in solution (50:1), intensity by size showed a decrease in the average size of large aggregates in solution.

The substantial intensity observed at the lower end of the log scale at ~ 10 nm for all the peptide-protein mixes trialed consistent with the control protein sample was attributed to free protein toroids. For peptide only samples (Figure 5.14 B), a narrow distribution of intensity by size at ~ 1000 nm for the highest concentration of peptide in solution suggested mostly peptide nanostructures that moved in solution in a similar manner.

To examine the molecular weight distribution of such nanostructures in solution, the peptide-protein mixes were subjected to sedimentation velocity analytical ultracentrifugation. As Prx3 peptide does not comprise of any aromatic amino acids, the relative absorbance in the peptide-protein mixes is contributed solely by the protein.

The plot of relative concentration vs sedimentation coefficient for the peptide-protein mixes revealed peak at ~ 8 S is consistent with the control protein sample though with lower relative absorbance (Figure 5.15). The result suggested that all the peptide-protein mixes observed of free protein toroids. The relative lower absorbance of the protein in solution for the peptide-protein mixes examined suggested that there was less free protein in solution.

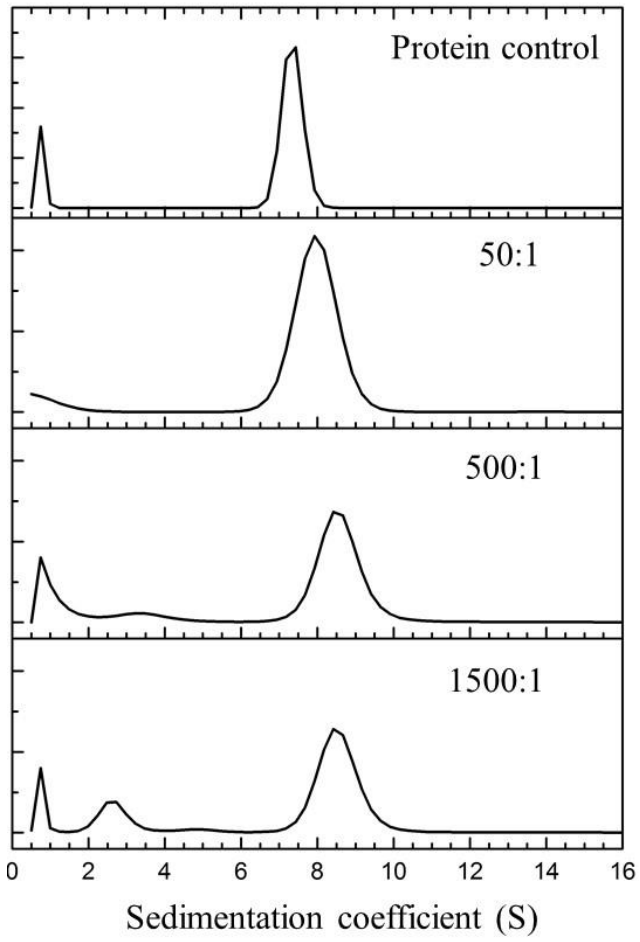


Figure 5.15: *Analytical ultracentrifugation of hPrx3 and Prx3 peptide + hPrx3 mixes in reducing pH 8.0 buffer. The final concentration of the peptide in the mix is: 0.50 mg/mL (1500:1); 0.17 mg/mL (500:1); 0.02 mg/mL (50:1) and protein control is hPrx3 at 0.1 mg/mL. The peptide-protein mix and control samples were incubated at 4°C for 24 hours prior to analysis.*

To assess the presence of extended nanotubular structures in peptide-protein mixes that may fall out of solution, the samples were analyzed by TEM. Negatively stained EM grids made from resuspended pellets of peptide-protein mixes containing Prx3 peptide at 1500 fold and 500 fold molar excess to protein revealed mostly extended nanotubular structures (Figure 5.16 A and B). A TEM of the peptide-protein mix containing the peptide at 50 fold molar excess to protein revealed a heterogeneous mix of short nanotubes and free toroids (Figure 5.16 C).

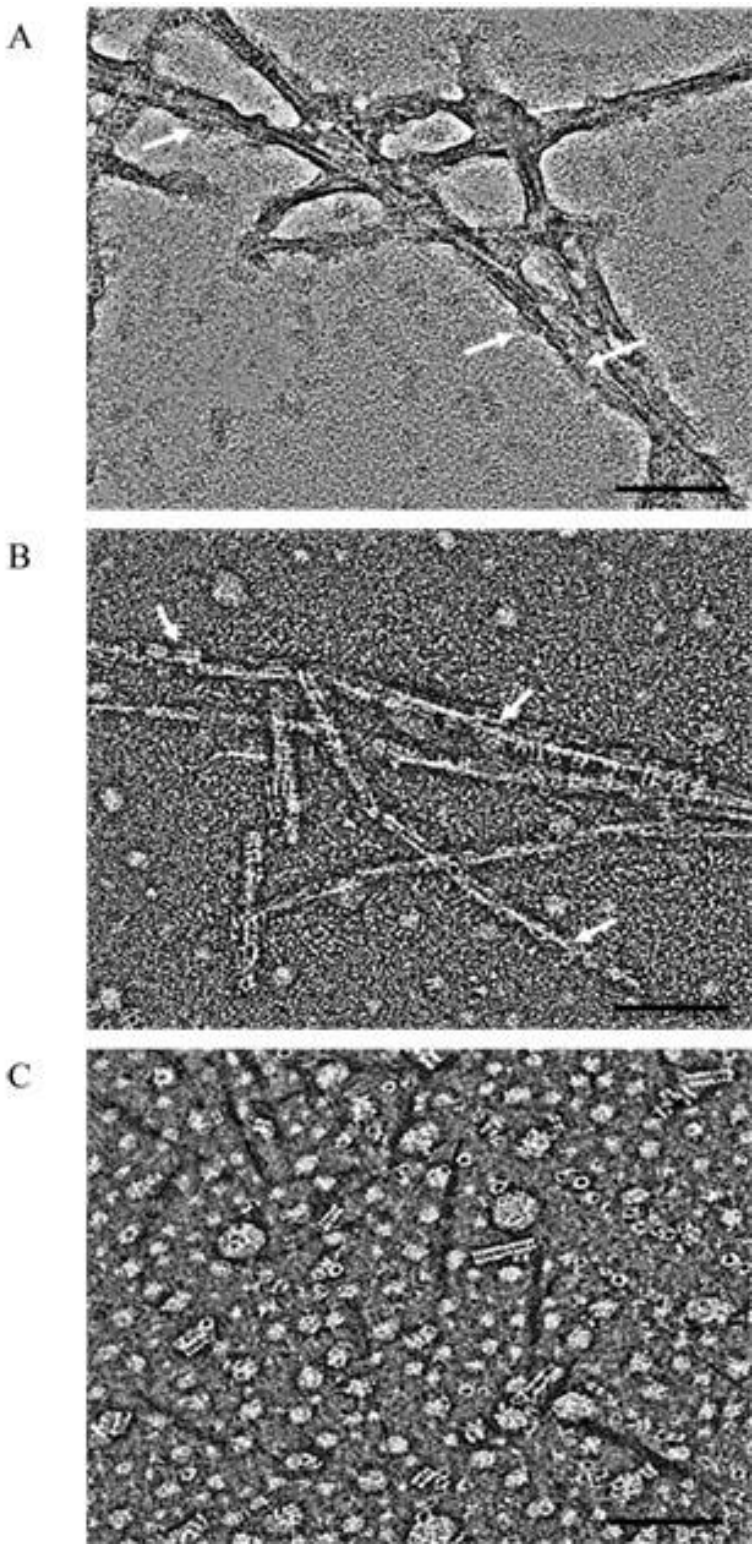


Figure 5.16: *TEM micrographs of pellets obtained post-centrifugation of peptide-protein mixes suspended in pH 8.0 buffer. (A) Peptide-protein mix at 1500:1; (B) Peptide-protein mix at 500:1; (C) Peptide-protein mix at 50:1. Protein concentration is 0.1 mg/mL. White arrows show the hPrx3 toroids laterally associated with extended peptide tapes. (Scale bar is 100 nm)*

The results suggested that the extended peptide nanostructures template *h*Prx3 protein toroids in solution and such large peptide-protein settle to the bottom of the tube, therefore not detected by AUC. For peptide-protein mix containing dilute amount of peptide mixed with protein, that contain few or no extended peptide nanostructures that can template protein toroids, only short toroidal stacks and free toroids are present. This result supports the inference that extended peptide nanostructures are necessary to template protein toroids in solution into extended nanotubes.

5.8 Discussion

In this work, self-assembling peptide sequence designed from the homodimer interface of Prx3 protein was successfully used to template nanotubular assembly of *h*Prx3 protein toroids. Peroxiredoxin proteins are versatile self-assembling protein tectons that can self-assemble into a range of different nanostructure in response to a physiological trigger. The protein tecton has been successfully exploited for the production of hierarchical assemblies such as nanostacks[36,38]. However, production of extended nanotubular assemblies of peroxiredoxin protein reaching microscale dimensions is reported here for the first time.

Peptide sequences that can self-assemble into extended nanostructures in solution[49] were used to template the higher order assembly of *h*Prx3 protein into nanotubes and nanosheets. Self-assembling peptides that form amyloid-like fibrils have been useful as scaffolds to link synthetic polymers[50,51] and biological molecules[52,24,25] to create nanostructures with functionality. The amphipathic peptide sequences described in chapter 3 and 4, all self-assembled in solution into extended nanostructures. The exposed side chains of charged and polar amino acid residues making up the peptide sequence provide multiple contact points for a large number of electrostatic interactions and hydrogen bonding. Charged biomolecule like

protein can interact and associate with such extended peptide structures to produce supramolecular complexes with distinct morphology.

To test this hypothesis, the toroid shaped *h*Prx3 protein was chosen due to the large surface area of the dodecamer quaternary conformation that is richly distributed with charged amino acid residues in a symmetrical manner. The self-assembling Prx3 peptide mixed with *h*Prx3 protein in buffer condition that favors the dodecamer conformation of the protein interestingly produced extended nanotubular structures of nano- and microscale lengths. The extended nanotubular assembly of *h*Prx3 protein with peptide nanotapes was consistent for both the tagged and untagged form of the protein. It is interesting that the N-terminal peptide tag attached to the protein to facilitate protein purification is known to favor the dodecamer oligomer form of the Prx3 protein in solution[36,39]. Also, recombinant constructs of peroxiredoxin containing the peptide tag have been shown to form short nanostacks templated through nickel functionalized gold nanoparticles[41]. However, the tag was not necessary for the specific extended nanotubular assembly of the *h*Prx3 protein with peptide nanotapes.

Experiments performed by mixing the peptide nanotapes with different toroid shaped proteins, revealed that the assembly was specific to the *h*Prx3 protein. Experiments with different self-assembling peptide sequences mixed with *h*Prx3 protein, however, resulted in extended nanotubular assembly in each case. The specific assembly of the protein toroids with the peptide can be explained by the large size of the dodecamer protein toroid which provides a large surface area of contact necessary for the assembly and its natural tendency to self-assemble. This is well evidenced in the self-assembly of protein toroids into stacks under acidic pH[11,36,38] and under crowded environments within the protein crystal[40].

As described in chapter 3 and chapter 4, the self-assembled peptide nanostructures dissociate upon dilution[49], therefore mixing experiments performed with dilute peptide acetate solution mixed with protein toroids revealed nanotubes that were smaller. Furthermore, mixing studies carried out with very dilute peptide solution mixed with protein toroids revealed mostly short nanotubes, nanostacks and free toroids. This suggested that the extent of the nanotube assembly upon mixing peptide with protein is relative and dependent on the amount and the length of peptide nanostructures in solution.

Solution characterization of peptide-protein mixes by DLS showed that the extended nanotubular assemblies were present in solution, though it was difficult to detect such supramolecular structures by AUC. Solution environment can have a substantial effect on molecular self-assembly processes, especially assembly of two different macromolecules through excluded volume effects, binding, and templating[53,54,55]. The absence of multiple peaks corresponding to differently sedimenting nanostacks and nanotubes for peptide-protein mixes examined by AUC suggested that under such dilute conditions the specific assembly is weakened. This can be attributed to the large solution partitioning that exists between the peptide aggregates and the protein toroids which hinder in molecular interactions necessary for the specific self-assembly. Protein toroids associated with the peptide nanostructures, however, quickly settle to the bottom of the solution and thus, not detected by AUC. This explains the low relative absorbance of free protein toroids for the peptide-protein mixes examined by AUC that corresponded to mostly free protein in solution.

The extended nanotubular structures observed for dilute peptide-protein mixes by TEM can be explained by molecular crowding that occurs on air-drying of the samples mounted onto the EM grid. This crowding of molecules results in the protein toroids and the peptide nanotapes

coming into close proximity which favors macromolecular associations that leads to templating of toroids onto peptide nanotapes.

Similarly, peptide-protein mixes with a large molar excess of peptide mixed with protein, presents a highly crowded environment with macromolecules in close contacts[53]. The higher number of diffusing molecules at higher concentration of peptide mixed with protein toroids results in more molecular collisions between the participating molecules leading to the increase in the tendency to form extended toroidal stacks. Also, it is important to note that at higher concentration in solution; the peptide forms viscous hydrogels comprised of self-assembled extended peptide tapes. Therefore, mixing higher fold excess of peptide in solution with protein toroids results in the increase in macroscopic viscosity of the final solution mix due to macromolecular crowding[56]. This change in viscosity results in slow translational movement of the macromolecules and the entropic effects under such crowded environments favor macromolecular association. This type of self-assembly of macromolecules in solution has been previously described in the case of proteins[17,57] and biopolymers[58].

Crystallization of proteins presents a similar crowded environment wherein protein units tend to self-assemble into hierarchical structures otherwise not observed in solution[55,56,57,58,59,60]. This explains why under highly concentrated conditions the protein toroids associated to the extended peptide nanostructures electrostatically, and the scaffolded toroids further lock with adjacent toroids through the R-interface[38] to form the stable extended nanotubular structures. The different assembly of protein toroids observed with the variant peptide H3Y can be explained by the templating of toroids onto the wide peptide tapes through the different exposed surface of the protein toroid.

The strategy described in this work proposes a single step process to template the assembly of *h*Prx3 protein into extended nanotubes and nanosheets in a controlled manner using self-assembling peptide sequences. Furthermore, the findings of this work show an alternate route of the assembly of peroxiredoxin toroids into extended tubular assemblies as observed in crowded cellular environment [11,35,61].

5.9 References

1. Seeman, N. C. (2010). Nanomaterials based on DNA. *Annual Review of Biochemistry* **79**, 65-87.
2. Rangnekar, A. & LaBean, T. H. (2014). Building DNA nanostructures for molecular computation, templated assembly, and biological applications. *ACS Accounts of Chemical Research* **47**, 1778-1788.
3. Endo, M., Takeuchi, Y., Emura, T., Hidaka, K. & Sugiyama, H. (2014). Preparation of chemically modified RNA origami nanostructures. *Chemistry* **20**, 15330-15333.
4. Li, H., Lee, T., Dziubla, T., Pi, F., Guo, S., Xu, J., Li, C., Haque, F., Liang, X.-J. & Guo, P. (2015). RNA as a stable polymer to build controllable and defined nanostructures for material and biomedical applications. *Nano Today* **10**, 631-655.
5. Woolfson, D. N. & Mahmoud, Z. N. (2010). More than just bare scaffolds: towards multi-component and decorated fibrous biomaterials. *Chemical Society Reviews* **39**, 3464-3679.
6. Brodin, J. D., Ambroggio, X. I., Tang, C., Parent, K. N., Baker, T. S. & Tezcan, F. A. (2012). Metal-directed, chemically tunable assembly of one-, two- and three-dimensional crystalline protein arrays. *Nature Chemistry* **4**, 375-382.
7. Gerrard, J. A. (2013). Protein nanotechnology: protocols, instrumentation, and applications. *Humana Press*.
8. Fromm, S. A., Bharat, T. A., Jakobi, A. J., Hagen, W. J. & Sachse, C. (2015). Seeing tobacco mosaic virus through direct electron detectors. *Journal of Structural Biology* **189**, 87-97.
9. Heddle, J. G. (2008). Protein cages, rings and tubes: useful components of future nanodevices? *Nanotechnology, Science and Applications* **1**, 67-78.
10. Lodish H, B. A. a. Z. S. (2000). Microtubule structures. In molecular cell biology. *W. H. Freeman*.
11. Saccoccia, F., Di Micco, P., Boumis, G., Brunori, M., Koutris, I., Miele, A. E., Morea, V., Sriratana, P., Williams, D. L., Bellelli, A. & Angelucci, F. (2012). Moonlighting by different stressors: crystal structure of the chaperone species of a 2-Cys peroxiredoxin. *Structure* **20**, 429-439.
12. Fletcher, D. A. & Mullins, R. D. (2010). Cell mechanics and the cytoskeleton. *Nature* **463**, 485-492.

13. Stradner, A., Sedgwick, H., Cardinaux, F., Poon, W. C. K., Egelhaaf, S. U. & Schurtenberger, P. (2004). Equilibrium cluster formation in concentrated protein solutions and colloids. *Nature* **432**, 492-495.
14. Krebs, M. R., Bromley, E. H., Rogers, S. S. & Donald, A. M. (2005). The mechanism of amyloid spherulite formation by bovine insulin. *Biophysical Journal* **88**, 2013-2021.
15. Grudzielanek, S., Velkova, A., Shukla, A., Smirnovas, V., Tatarek-Nossol, M., Rehage, H., Kapurniotu, A. & Winter, R. (2007). Cytotoxicity of insulin within its self-assembly and amyloidogenic pathways. *Journal of Molecular Biology* **370**, 372-384.
16. Yan, H., Frielinghaus, H., Nykanen, A., Ruokolainen, J., Saiani, A. & Miller, A. F. (2008). Thermoreversible lysozyme hydrogels: properties and an insight into the gelation pathway. *Soft Matter* **4**, 1313-1325.
17. Pan, W., Filobelo, L., Pham, N. D., Galkin, O., Uzunova, V. V. & Vekilov, P. G. (2009). Viscoelasticity in homogeneous protein solutions. *Physical Review Letters* **102**, 058101-058104.
18. Javid, N., Vogtt, K., Roy, S., Hirst, A. R., Hoell, A., Hamley, I. W., Ulijn, R. V. & Sefcik, J. (2011). Supramolecular structures of enzyme clusters. *Journal of Physical Chemistry Letters* **2**, 1395-1399.
19. Nguyen, A. T., Huang, Q. L., Yang, Z., Lin, N., Xu, G. & Liu, X. Y. (2015). Crystal networks in silk fibrous materials: from hierarchical structure to ultra performance. *Small* **11**, 1039-1054.
20. Gerrard, J. A. (2013). Protein nanotechnology: what is it? in protein nanotechnology: protocols, instrumentation, and applications. *Humana Press.*
21. Sobti, M. & Mabbutt, B. C. (2013). Rational-based protein engineering: tips and tools in protein nanotechnology: protocols, instrumentation, and applications. *Humana Press*.
22. Collins, B. M., Harrop, S. J., Kornfeld, G. D., Dawes, I. W., Curmi, P. M. & Mabbutt, B. C. (2001). Crystal structure of a heptameric Sm-like protein complex from archaea: implications for the structure and evolution of snRNPs. *Journal of Molecular Biology* **309**, 915-923.
23. Aldeek, F., Safi, M., Zhan, N., Palui, G. & Mattoussi, H. (2013). Understanding the self-assembly of proteins onto gold nanoparticles and quantum dots driven by metal-histidine coordination. *ACS Nano* **7**, 10197-10210.

24. Rożkiewicz, D. I., Myers, B. D. & Stupp, S. I. (2011). Interfacial self-assembly of cell-like filamentous microcapsules. *Angewandte Chemie International Edition in English* **50**, 6324-6327.
25. Kuang, Y., Yuan, D., Zhang, Y., Kao, A., Du, X. & Xu, B. (2013). Interactions between cellular proteins and morphologically different nanoscale aggregates of small molecules. *RSC Advances* **3**, 7704-7707.
26. Hudalla, G. A., Sun, T., Gasiorowski, J. Z., Han, H., Tian, Y. F., Chong, A. S. & Collier, J. H. (2014). Gradated assembly of multiple proteins into supramolecular nanomaterials. *Nature Materials* **13**, 829-836.
27. Wood, Z. A., Poole, L. B., Hantgan, R. R. & Karplus, P. A. (2002). Dimers to doughnuts: redox-sensitive oligomerization of 2-cysteine peroxiredoxins. *Biochemistry* **41**, 5493-5504.
28. Flohé, L. & Harris, J. R. (2007). Peroxiredoxin systems: structures and functions, *Springer Science*.
29. Jang, H. H., Lee, K. O., Chi, Y. H., Jung, B. G., Park, S. K., Park, J. H., Lee, J. R., Lee, S. S., Moon, J. C., Yun, J. W., Choi, Y. O., Kim, W. Y., Kang, J. S., Cheong, G. W., Yun, D. J., Rhee, S. G., Cho, M. J. & Lee, S. Y. (2004). Two enzymes in one; two yeast peroxiredoxins display oxidative stress-dependent switching from a peroxidase to a molecular chaperone function. *Cell* **117**, 625-635.
30. Meissner, U., Schroder, E., Scheffler, D., Martin, A. G. & Harris, J. R. (2007). Formation, TEM study and three-dimensional reconstruction of the human erythrocyte peroxiredoxin-2 dodecahedral higher-order assembly. *Micron* **38**, 29-39.
31. Cao, Z., Roszak, A. W., Gourlay, L. J., Lindsay, J. G. & Isaacs, N. W. (2005). Bovine mitochondrial peroxiredoxin III forms a two-ring catenane. *Structure* **13**, 1661-1664.
32. Phalen, T. J., Weirather, K., Deming, P. B., Anathy, V., Howe, A. K., van der Vliet, A., Jonsson, T. J., Poole, L. B. & Heintz, N. H. (2006). Oxidation state governs structural transitions in peroxiredoxin II that correlate with cell cycle arrest and recovery. *Journal of Cell Biology* **175**, 779-89.
33. Gourlay, L. J., Bhella, D., Kelly, S. M., Price, N. C. & Lindsay, J. G. (2003). Structure-function analysis of recombinant substrate protein 22 kDa (SP-22). A mitochondrial 2-cys peroxiredoxin organized as a decameric toroid. *Journal of Biological Chemistry* **278**, 32631-32637.

34. Harris, J. R., Schröder, E., Isupov, M. N., Scheffler, D., Kristensen, P., Littlechild, J. A., Vagin, A. A. & Meissner, U. (2001). Comparison of the decameric structure of peroxiredoxin-II by transmission electron microscopy and X-ray crystallography. *Biochimica et Biophysica Acta: Protein Structure and Molecular Enzymology* **1547**, 221-234.

35. Kato, H., Asanoi, M., Nakazawa, T. & Maruyama, K. (1985). Cylinder protein isolated from rat liver mitochondria. *Zoological Science* **2**, 485-490.

36. Angelucci, F., Saccoccia, F., Ardini, M., Boumis, G., Brunori, M., Di Leandro, L., Ippoliti, R., Miele, A. E., Natoli, G., Scotti, S. & Bellelli, A. (2013). Switching between the alternative structures and functions of a 2-Cys Peroxiredoxin, by site-directed Mutagenesis. *Journal of Molecular Biology* **425**, 4556-4568.

37. Phillips, A. J., Littlejohn, J., Yewdall, N. A., Zhu, T., Valery, C., Pearce, F. G., Mitra, A. K., Radjainia, M. & Gerrard, J. A. (2014). Peroxiredoxin is a versatile self-assembling tecton for protein nanotechnology. *Biomacromolecules* **15**, 1871-1881.

38. Cao, Z., McGow, D. P., Shepherd, C. & Lindsay, J. G. (2015). Improved catenated structures of bovine peroxiredoxin III F190L reveal details of ring-ring Interactions and a novel conformational state. *PLoS ONE* **10**, 1-15.

39. Radjainia, M., Venugopal, H., Desfosses, A., Phillips, Amy J., Yewdall, N. A., Hampton, M. B., Gerrard, J. A. & Mitra, A. K. Cryo-electron microscopy structure of human peroxiredoxin-3 filament reveals the assembly of a putative chaperone. *Structure* **23**, 912-920.

40. Gretes, M. C. & Karplus, P. A. (2013). Observed octameric assembly of a plasmodium yoelii peroxiredoxin can be explained by the replacement of native "ball-and-socket" interacting residues by an affinity tag. *Protein Science* **22**, 1445-1452.

41. Yewdall, N. A., Venugopal, H., Desfosses, A., Abrishami, V., Yosaatmadja, Y., Hampton, Mark B., Gerrard, Juliet A., Goldstone, David C., Mitra, Alok K. & Radjainia, M. Structures of human peroxiredoxin 3 suggest self-chaperoning sssembly that maintains catalytic state. *Structure* **24**, 1120-1129.

42. Ardini, M., Giansanti, F., Di Leandro, L., Pitari, G., Cimini, A., Ottaviano, L., Donarelli, M., Santucci, S., Angelucci, F. & Ippoliti, R. (2014). Metal-induced self-assembly of peroxiredoxin as a tool for sorting ultrasmall gold nanoparticles into one-dimensional clusters. *Nanoscale* **6**, 8052-8061.

43. Saccoccia, F., Di Micco, P., Boumis, G., Brunori, M., Koutris, I., Miele, A. E., Morea, V., Sriratana, P., Williams, D. L., Bellelli, A. & Angelucci, F. (2012). Moonlighting by different stressors: crystal structure of the chaperone species of a 2-Cys peroxiredoxin. *Structure* **20**, 429-439.
44. Orlova, E. V. & Saibil, H. R. (2011). Structural analysis of macromolecular assemblies by electron microscopy. *ACS Chemical Reviews* **111**, 7710-7748.
45. Nakamura, T., Yamamoto, T., Inoue, T., Matsumura, H., Kobayashi, A., Hagihara, Y., Uegaki, K., Ataka, M., Kai, Y. & Ishikawa, K. (2006). Crystal structure of thioredoxin peroxidase from aerobic hyperthermophilic archaeon Aeropyrum pernix K1. *Proteins: Structure, Function, and Bioinformatics* **62**, 822-826.
46. Ferré-D'Amaré, A. R. & Burley, S. K. (1997). Dynamic light scattering in evaluating crystallizability of macromolecules. In methods in enzymology.*Academic Press*.
47. Lorber, B., Fischer, F., Bailly, M., Roy, H. & Kern, D. (2012). Protein analysis by dynamic light scattering: methods and techniques for students. *Biochemistry and Molecular Biology Education* **40**, 372-382.
48. Filipe, V., Hawe, A. & Jiskoot, W. (2010). Critical evaluation of nanoparticle tracking analysis (NTA) by nanosight for the measurement of nanoparticles and protein aggregates. *Pharmaceutical Research* **27**, 796-810.
49. Valery, C., Pandey, R. & Gerrard, J. A. (2013). Protein β-interfaces as a generic source of native peptide tectons. *RSC Chemical Communications* **49**, 2825-2827.
50. Krishna, O. D. & Kiick, K. L. (2010). Protein- and peptide-modified synthetic polymeric biomaterials. *Biopolymers* **94**, 32-48.
51. Hardy, J. G., Amend, M. N., Geissler, S., Lynch, V. M. & Schmidt, C. E. (2015). Peptide-directed assembly of functional supramolecular polymers for biomedical applications: electroactive molecular tongue-twisters (oligoalanine-oligoaniline-oligoalanine) for electrochemically enhanced drug delivery. *Journal of Materials Chemistry B* **3**, 5005-5009.
52. Cui, H., Webber, M. J. & Stupp, S. I. (2010). Self-Assembly of peptide amphiphiles: from molecules to nanostructures to biomaterials. *Biopolymers* **94**, 1-18.
53. Zimmerman, S. B. & Minton, A. P. (1993). Macromolecular crowding: biochemical, biophysical, and physiological consequences. *Annual Review of Biophysics and Biomolecular Structure* **22**, 27-65.
54. Ellis, R. J. (2001). Macromolecular crowding: obvious but underappreciated. *Trends in Biochemical Science* **26**, 597-604.

55. Minton, A. P. (2000). Implications of macromolecular crowding for protein assembly. *Current Opinion in Structural Biology* **10**, 34-39.

56. Wieczorek, G. & Zielenkiewicz, P. (2008). Influence of macromolecular crowding on protein-protein association rates: a brownian dynamics study. *Biophysical Journal* **95**, 5030-5036.

57. Hu, Z., Jiang, J. & Rajagopalan, R. Effects of macromolecular crowding on biochemical reaction equilibria: a molecular thermodynamic perspective. *Biophysical Journal* **93**, 1464-1473.

58. Groen, J., Foschepoth, D., te Brinke, E., Boersma, A. J., Imamura, H., Rivas, G., Heus, H. A. & Huck, W. T. S. (2015). Associative interactions in crowded solutions of biopolymers counteract depletion effects. *Journal of the American Chemical Society* **137**, 13041-13048.

59. Kim, J. S. & Yethiraj, A. (2009). Effect of macromolecular crowding on reaction Rates: a computational and theoretical study. *Biophysical Journal* **96**, 1333-1340.

60. Weyand, S., Kefala, G., Svergun, D. I. & Weiss, M. S. (2009). The three-dimensional structure of diaminopimelate decarboxylase from mycobacterium tuberculosis reveals a tetrameric enzyme organisation. *Journal of Structural and Functional Genomics* **10**, 209-217.

61. Jang, H. H., Lee, K. O., Chi, Y. H., Jung, B. G., Park, S. K., Park, J. H., Lee, J. R., Lee, S. S., Moon, J. C., Yun, J. W., Choi, Y. O., Kim, W. Y., Kang, J. S., Cheong, G.-W., Yun, D.-J., Rhee, S. G., Cho, M. J. & Lee, S. Y. (2004). Two enzymes in one: two yeast peroxiredoxins display oxidative stress-dependent switching from a peroxidase to a molecular chaperone function. *Cell* **117**, 625-635.

6 Chapter Six

Self-assembly of helical coiled coil interface derived peptides

6.1 Introduction

Helical coiled coils are another important protein motif that has been explored for the design of self-assembling peptide tectons[1]. α-helical coiled coils are versatile protein-protein interactive domains that constitute nearly 3% of reported protein coding regions[2]. They were first identified in fibrous assemblies of proteins such as keratin, myosin, epidermin, fibrinogen and soft α-keratin (stratum corneum) and hard α-keratin (trichocytes)[3,4]. Helical coiled coils are comprised of two or more helices wrapped around each other. This structural motif plays an important role in cementing protein-protein interactions for oligomerization of proteins involved in diverse biological processes[1]. The hallmark of such helical coiled coils is the repetitive arrangement of seven residues designated "abcdefg" called the heptad motif[5,6]. Sequences designed from heptad motifs found in natural assemblies of different structural proteins such as collagen, albumin, gelatin, fibroin and the intermediate filament protein keratin have been investigated for development of biomaterials[7,8]. Permutations and combinations of native heptad motifs have allowed for production of peptide tectons that can be triggered to self-assemble into a range of different fibrous nanostructures[1].

Keratins show much promise for design of stable biomaterials due to their inherent property to self-assemble into mechanically durable, highly stable and tough fibrous structures that form the bulk of epidermal appendages such as hair, wool, horns, hooves and nails, as well as their natural abundance[9]. Keratins have been used to produce biomaterials that are biocompatible and biodegradable. Keratins from wool and human hair have specific cell binding motifs with

the sequence LDV and EDS that support cellular attachment. Also, keratins are believed to regulate cellular functions and behavior[11]. These properties has allowed the use of keratin based materials for 3D cell culture and for tissue engineering applications. Self-assembled keratin films produced by mixing with glycerol and chitosan exhibit enhanced mechanical properties and have been successfully used to produce films with wound healing and antibacterial activity[10]. Another important example is the keratin-polyamide 6 (PA6) nanofibers produced by electrospinning of solution mixes containing different rations of keratin protein and polyamide 6 (PA6). Keratin-PA6 films and nanofibers have been used in variety of applications ranging from biomedical sciences, for active water filtration and as textile fibers. Keratin based hydrogels demonstrate neuroinductive ability and facilitate tissue regeneration *in vitro* through enhanced activity of Schwann cells, and have been very useful as biomaterials in regenerative medicine[11,12].

Mammalian keratins are versatile proteins produced in vertebrate epithelia that rapidly self-assemble in a heteromeric manner into smooth and flexible filaments[13,14,15,16] together with keratin associated proteins that are intermediate in diameter (~ 10 nm)[17,18]. Keratin IF proteins that make up the hard tissues such as hair and nail ("hard" keratins) and the keratin IF of various soft epithelia ("soft" keratin) are made of low sulfur polypeptides which are homologous: low molecular weight (~ 45 kDa) acidic type I and the higher molecular weight (~ 60 kDa) neutral to basic type II keratins[19,20]. Hard keratin IFs make up a large portion of the skin epidermis and all the dry weight in keratinized epidermis such as wool, hair, nail, claws, quills, horns and hooves[21]. Keratin fibers are associated with cortical cells comprised of longitudinally aligned intermediate filaments in the growth direction surrounded by cell matrix. Figure 6.1 shows a schematic of the keratin fiber[19].

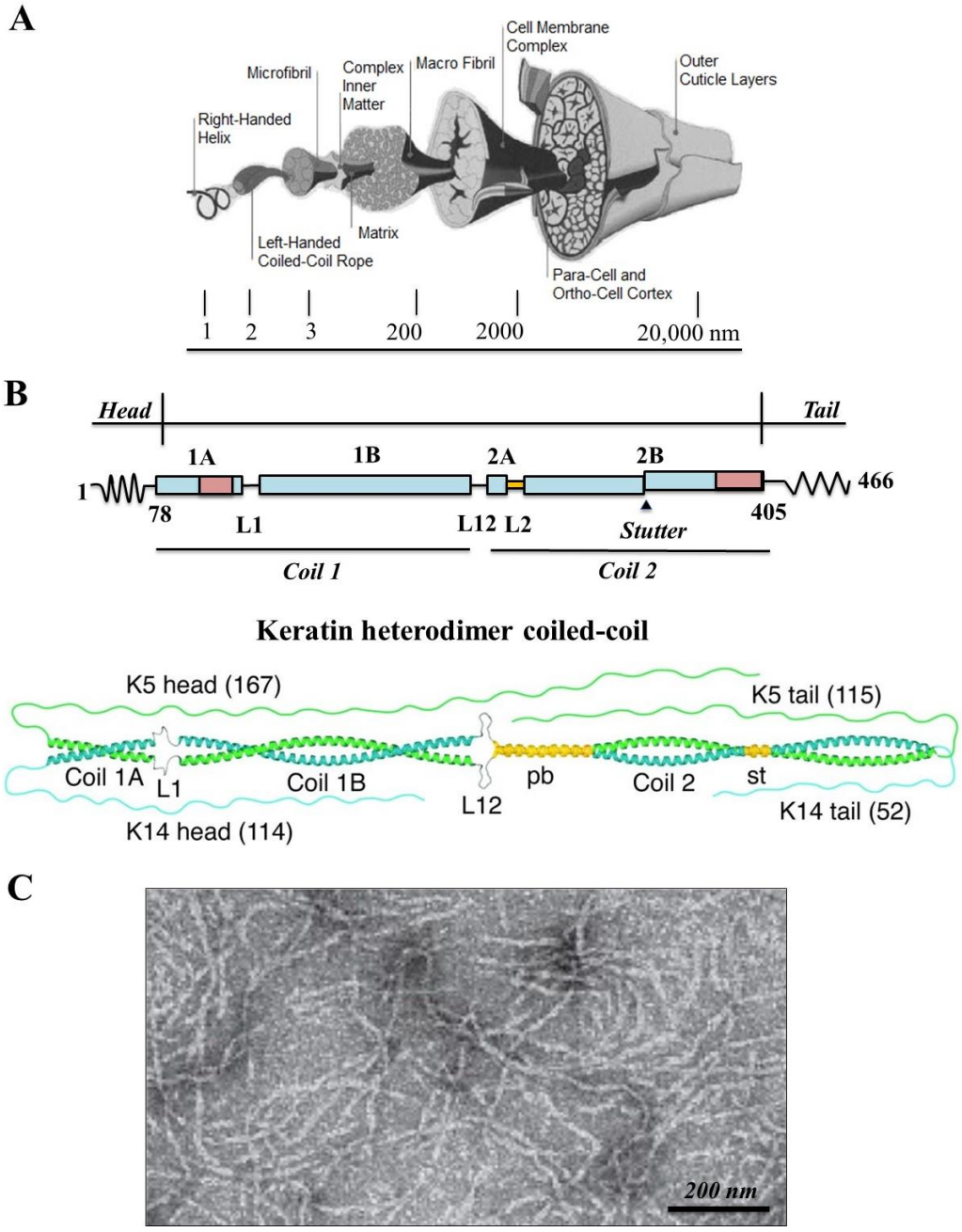


Figure 6.1: *Keratin intermediate filament organization: (A) Schematic diagram showing the structural organization of the intermediate filament protein keratin in the wool fiber ;(B) Model of IF protein structure showing the central helical coiled coil rod domain separated by linking segments and the coiled coil heterodimer complex formed from the self-assembly of type I and type II keratins; (C) TEM micrographs showing the morphology of keratin intermediate filaments. (Taken with permission from CSIRO[22] and Hermann et al. 1999[23].*

Most subtypes of IF proteins self-assemble as a homopolymer, but keratin intermediate filaments that form wool fibers are distinct in that they assemble into an obligate heteropolymer. Self-assembly of hard α-keratins into IFs proceeds from dimerization of type I and type II keratins into a coiled coil heterodimer[23]. Oligomerization of coiled coil heterodimers gives rise to the second level of structural hierarchy the unit length filaments (ULFs) with width between 50–60 nm or 50–120 nm. Annealing of the ULFs leads to elongation into long extended filaments, which further undergo compaction along the lateral axis to give rise to ~10 nm wide mature intermediate filaments that make wool fibers[23,24].

New Zealand (NZ) is one of the largest producers of wool and NZ wool exports contribute 11% of the annual global wool production[25]. The value of the wool produced is dependent on the quality of the wool, which is in turn dependent on the quality of the individual fibers. To open up new, non-traditional, markets for NZ wool, wool basic components can be explored as potential sources of new biomaterials for a range of applications, including medical and material sciences. To do so, the basic structure and properties of intermediate filament proteins need to be understood at molecular level. In particular the self-assembly process leading to the formation of the 10 nm intermediate filaments (IFs) that act as basic building blocks in the hierarchical assembly of wool fibers[16].

This chapter discusses the design of peptide sequences from keratin proteins that are the main components of wool intermediate filaments and the basic self-assembling properties of chosen fragment peptide sequences investigated.

6.2 Keratin as source of peptide tectons

Crystallographic studies focused on the central rod domain of the intermediate filament proteins[18,26,27,28] revealed that specific motifs located in the C-terminal end of the helical rod domain are critical for the self-assembly of the heterodimers and subsequent filament assembly[29]. Such motifs show high consensus among different keratins across species as diverse as primitive fish and humans, which include the non-heptad sequence called "stutter" and the highly conserved so called consensus motifs: "LNDR" motif in coil 1A and "TYRKLLEGEE" motif in coil 2B[16,27,30]. These motifs are part of the coiled coil protein interface that is involved in the oligomerization of IF proteins.

In the crystal structure of IF protein vimentin homodimer coil 2B and keratin heterodimer coil 2B, the consensus motif sequence "TYRKLLEGEE" is found to be in complementary non-covalent close contact[27,31]. Structural analysis of vimentin homodimer coil 2B (PDB 1GK6) and keratin K5-K14 heterodimer coil 2B (PDB 3TNU) for interface sequences by the PISA software revealed the consensus sequence motif as an interface motif. Inspection of the interface residues revealed two important interactions: a salt bridge between the glutamic acid residue and arginine residue and a hydrogen bond between the tyrosine residue and arginine residue (Figure 6.2). The salt bridge between the glutamic acid residue and the arginine residue within the consensus sequence motif is an essential interaction across the interface and conserved in both homopolymeric and heteropolymeric IFs[32,33]. The structural versatility of such interface motifs provides the necessary driving force for such oligomerization of proteins through sets of non-covalent interactions. Also, such motifs are evolutionary optimized to drive such self-assembly in a highly specific and controlled manner[34].

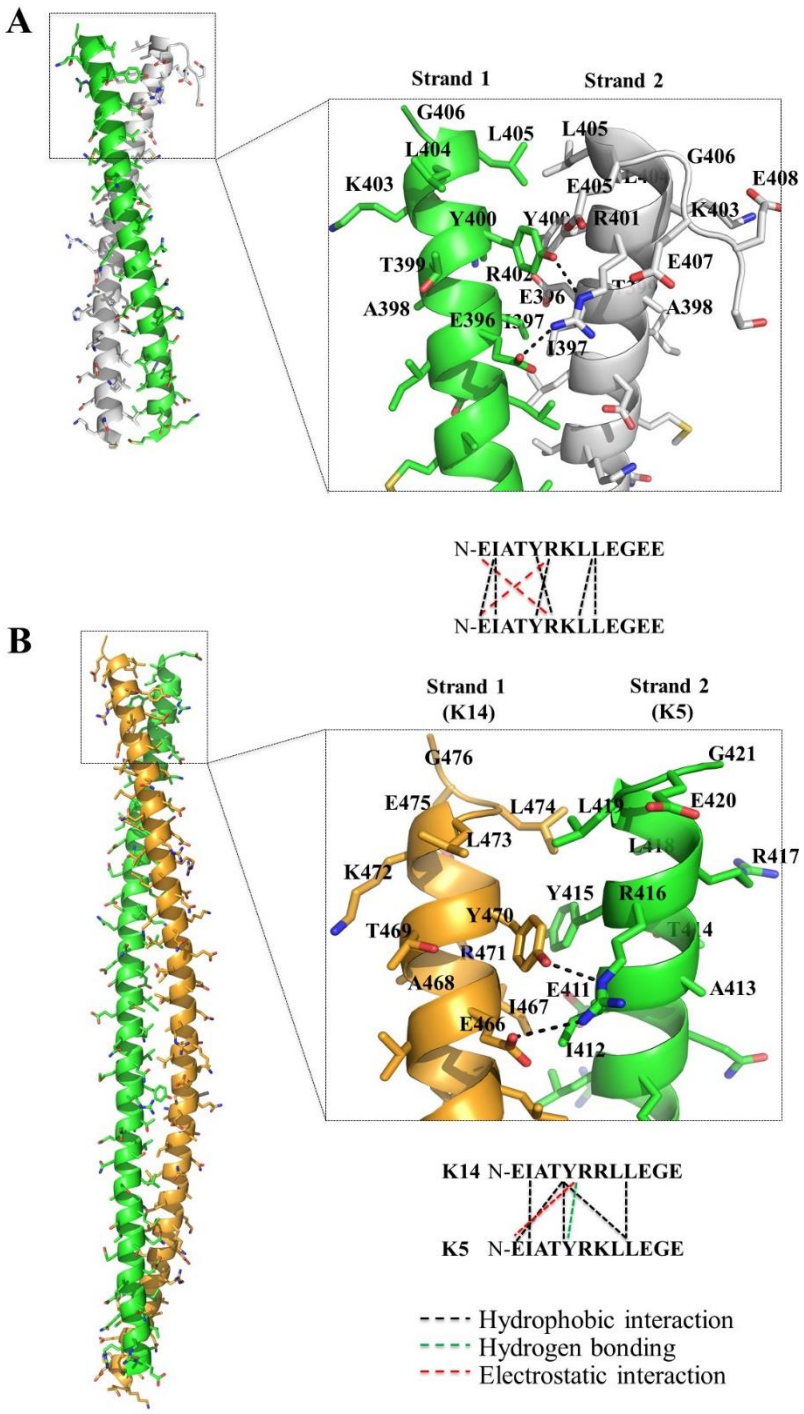


Figure 6.2: *Atomic structure of coil 2B domain: Crystal structure of vimentin homodimer (A, PDB 1GK6) and human keratin K5/K14 heterodimer (B; PDB 3TNU). The molecular interactions occurring between the consensus sequence motif at coiled coil protein interface of the vimentin homodimer and K5-K14 heterodimer are highlighted*[15,32].

Another important feature of the coil 2B of IF keratins is the rich distribution of cysteine residues, especially flanking the consensus sequence motif of IF keratins specific to the wool fiber[35]. Current understanding is that the disulfide bonding within keratin filaments is part of a major reorganization of the filaments from predominantly parallel bundles to crossed networks upregulated at early stages of terminal differentiation of epidermal cells[32,36,37]. Keratins extracted from the stratum corneum in the absence of reducing agents and in the presence of air rapidly disulfide cross-link into large aggregates unable to enter an SDS-PAGE gel[36]. In another study, mutant xenopus and mouse keratin designed with a cysteine residue introduced in the central rod domain formed disulfide linked keratin homodimers over time in the presence of air[38]. Sequence analysis of wool hard α-keratin type I and type II focused on the coiled coil 2B region generated using T-coffee and ESpript (http://espript.ibcp.fr)[39] revealed cysteine residue present flanking the consensus sequence motif "CEINTYRS/GLLE/DSEDC" (Figure 6.3). The dense localization of cysteine residues near the highly conserved consensus sequence motif at the C-terminus within hard α-keratins of human hair and sheep wool makes them sensitive to redox change.

In the specific case of wool hard α-keratins, the only heptad like sequence studied thus far looking at self-assembly is the consensus sequence motif "YRRLLEGEE". This motif was identified as a key coil 2B motif that can self-assemble into tetramers on its own[27,40]. In an earlier study, Hatzfeld *et al*. showed that a synthetic peptide corresponding to the consensus sequence motif of human cytoskeletal keratin containing an N-terminal cysteine residue inhibits keratin filament assembly. However, any specific self-assembly of the peptide sequence on its own was not described[41].

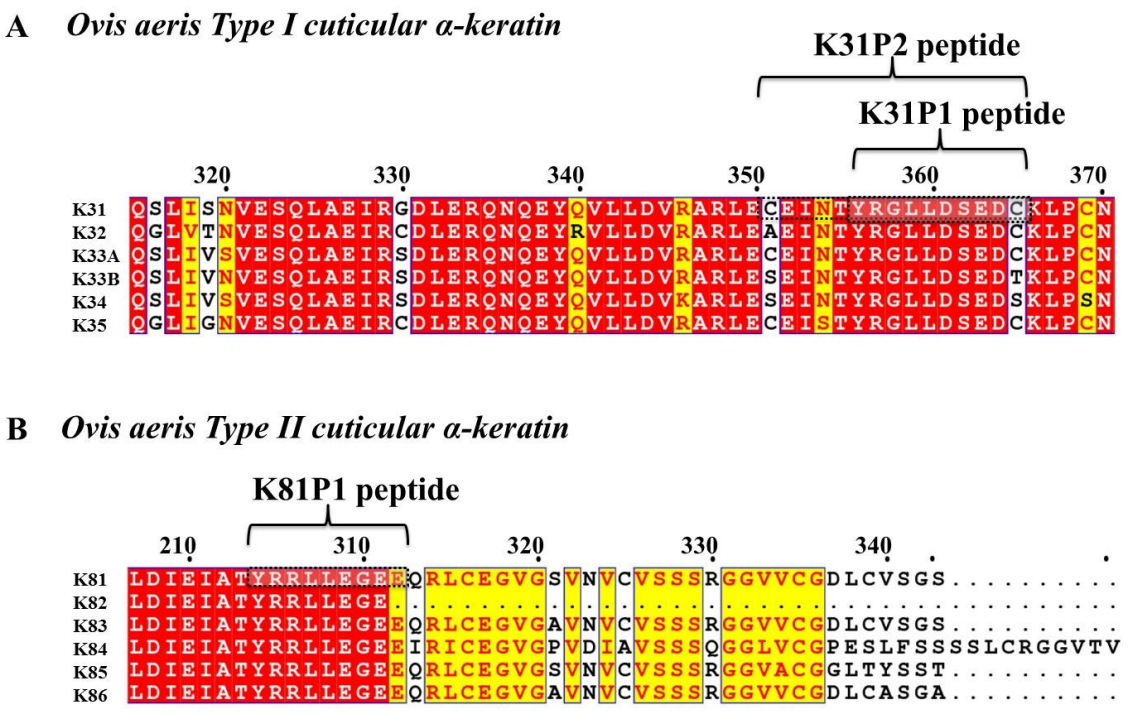


Figure 6.3: *Sequence alignment of C-terminus end of the coiled coil 2B region of the central rod domain of different type I and type II hard α-keratins from Ovis aeris. The black dashed box shows the location of the peptide sequence comprised of the consensus sequence motif designed from type I α-keratin (K31P1 and K31P2) and type II α-keratin (K81P1).*

To our knowledge, no other isolated wool peptide motif has been investigated as source of self-assembling peptides. The influence of cysteine residues, despite their intrinsic role in intermediate filament cross-linking and responsiveness to redox conditions, has also not been previously evaluated in terms of influence and control of self-assembly.

6.3 Peptide sequences investigated for self-assembly

A UniProt wide search for cuticular α-keratins and keratin associated proteins in *Ovis aeris* species resulted in > 240 entries. The entries were narrowed down by focusing on: (i) type I and type II keratins, while excluding keratin associated proteins; (ii) the consensus motif

sequence at the carboxy-terminal end of coil 2B in the central coiled coil rod domain and (iii) the region of consensus sequence motif flanked by cysteine residues.

The final set of peptide sequences were composed of 11-15 residues corresponding to the consensus sequence motif from the C-terminus region of the 2B coil of *O. aeris* type I hard α-keratin and *O. aeris* type II hard α-keratin. The peptide sequences were modified from the original sequence by addition of the cysteine residue at both the N-terminus and C-terminus and capped at both the ends: N-terminal acetylated and C-terminal amidated to conserve the overall charge of the sequence (Table 6.1).

Table 6.1: *Ovine keratin derived peptide sequences chosen as models*

Name	Peptide	Source	pI	Charge at neutral pH
K31P1	Ac-CYRGLLDSEDC-NH_2	Keratin 31	3.54	-2.1
K81P1	Ac-CYRRLLEGEEC-NH_2	Keratin 81	4.45	-1.1
K31P2	Ac-CEINTYRGLLDSEDC-NH_2	Keratin 31	3.43	-3.1

Note: Figure 6.3 shows the location of the corresponding source peptide sequence within the type I and type II hard α-keratins.

In the specific case of the K31P2 peptide derived from the C-terminus end of the coiled coil 2B of type I hard α-keratin, it is noteworthy that the peptide sequence comprised of the consensus sequence motif flanked by cysteine residues is a naturally conserved sequence (Figure 6.3 A). The K31P2 peptide investigated in this work differs from the previously reported synthetic peptide sequence[41] in the total length, position and number of cysteine residues.

6.4 Characterization of peptide self-assembly

6.4.1 Solution characterization of peptide self-assembly

The peptide sequences were obtained as a white colored lyophilized powder from Mimotopes Ltd. (Sydney, Australia). The purity of the peptides was > 95% as tested by the manufacturer. Initial studies on the peptide acetate dissolved in water and HFIP/TFA mixes showed good solubility. However studies with peptide acetate dissolved in water were not soluble at peptide concentrations > 10 mg/mL (w/v). The secondary structure of the peptide sequences in solution, was examined by circular dichroism spectra for dilute peptide acetates in water at 0.1 mg/mL. CD spectra for all the three peptide showed minima at ~ 222 nm and at ~ 208 nm (Figure 6.4). The presence of the double minima at the corresponding wavelengths for each of the peptides suggested a helix conformation of peptides in solution[42].

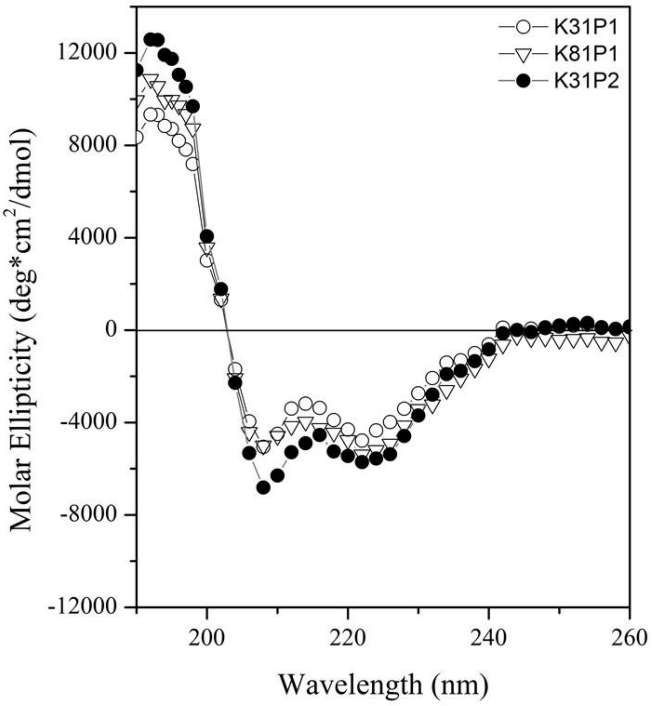


Figure 6.4: *Far UV circular dichroism spectra of peptide acetates K31P1, K81P1 and K31P2 in water at 0.1 mg/mL. The peptide acetate in water samples was equilibrated for 24 hours at 4°C and degassed prior to analysis.*

6.4.2 TEM characterization of peptide self-assembly

Lyophilized peptide acetates were dissolved in water in the 10–50 mg/mL (w/v) range. In the specific case epidermal keratins the assembly of keratins into extended intermediate filaments is known to occur at room temperature[43]. Thus, the experiments were carried out at room temperature and the peptide acetates were equilibrated in solution for at least 24 hours prior to analysis.

Peptide K31P1 was found to be soluble in water at all the concentrations investigated, however, both peptide K81P1 and peptide K31P2 showed lower solubility in water. Furthermore, peptide acetates K81P1 and K31P2 in water at ≥ 10 mg/mL concentration, stored at 25°C (room temperature) for over 24 hours showed visible precipitates (Figure 6.5). This indicated a saturated peptide solution having large number of insoluble peptide aggregates.

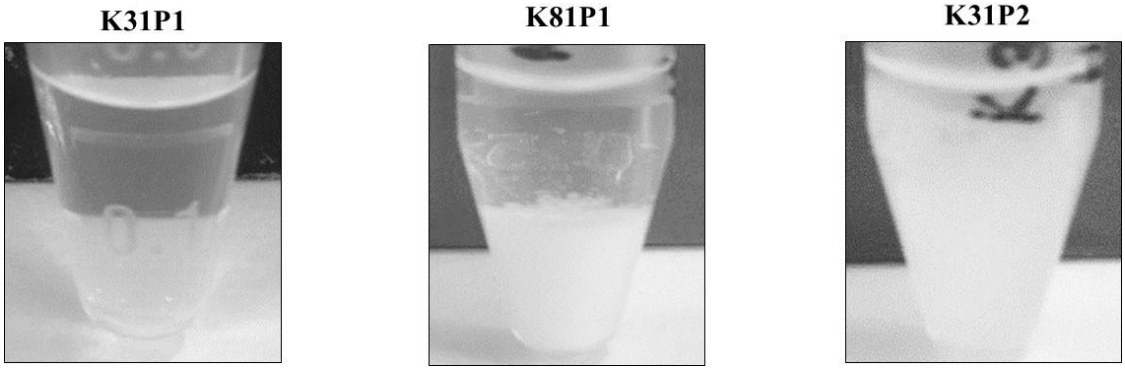


Figure 6.5: *Peptide acetates in water (pH ~ 5.0) at 10 mg/mL equilibrated for 24 hours at 25°C.*

Negatively stained micrographs of the peptide acetate in water samples for K31P1 and K81P1 analyzed by TEM (Section 2.9.1) did not show any substantial difference between EM micrographs prepared immediately after dilution and 24 hours after equilibration at 25°C. It is important to note that the CD spectra showed that the peptides attained a helical conformation

in water, however absence of any clear self-assembly by TEM suggested the peptide self-assembled into short helical coiled-coil structures which were difficult to visualize by TEM. A similar concentrated sample of K31P2 peptide equilibrated over 24 hours however, self-assembled into short tape like structures that appeared to be hierarchical structures self-assembled from smaller rod shaped structures (Figure 6.6).

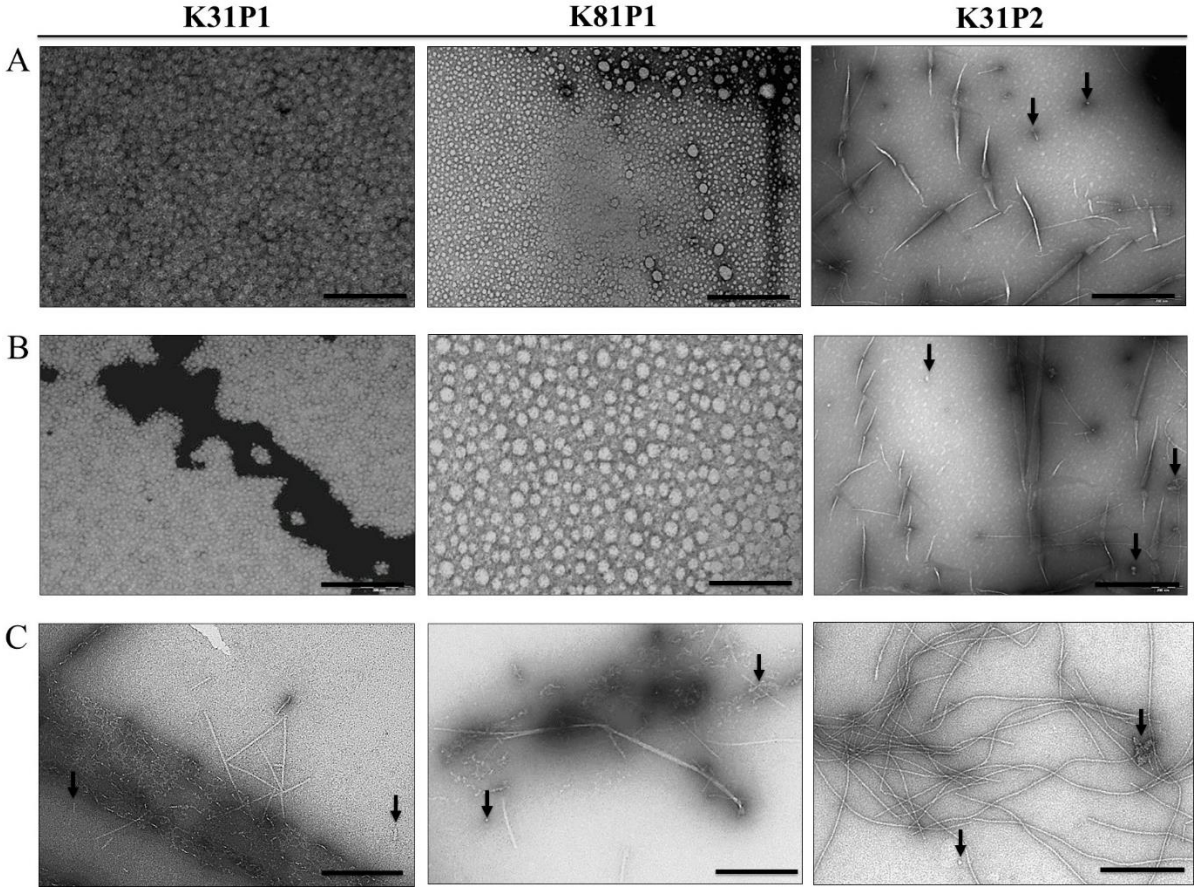


Figure 6.6: *Negatively stained micrographs of peptide acetates in water (pH ~ 5.0) at 10 mg/mL (w/v) analyzed at 24 hours (A), T1 week (B) and T3 weeks (C) stored at 25°C . Arrows indicate to the short rod shaped fragments. (Scale bar is 200 nm)*

Peptide acetates prepared in water at 10 mg/mL monitored over three weeks stored at room temperature developed a turbid appearance suggesting the presence of extended assemblies. Such samples of peptide acetates analyzed by TEM showed mostly short rod like fragments

and some longer tapes for the K31P1 peptide; and mostly extended tape like nanostructures for K81P1 peptide; extended filament like nanostructures and short rod shaped fragments for K31P2 peptide (Figure 6.6 C).

The presence of a heterogeneous mix of short rod like fragments and extended filament structures in all the three cases suggested that the peptide first self-assembled into short rod like fragments that further assembled into extended filament like nanostructures. The development of such extended assemblies in solution over time explains the turbid appearance of the peptide acetate solutions. Further, the absence of such rod like fragments and tapes for K31P1 and K81P1 peptide acetate samples equilibrated for 24 hours and over 1 week suggested a slow assembly process for these peptide sequences.

Measurements performed on the extended nanostructures observed by TEM for the peptide acetates using ImageJ software, revealed nanostructures of mostly 50 nm in length for K31P1 peptide, 200 nm in length for K81P1 peptide and between 400 nm and 800 nm in length for K31P2 peptide (Figure 6.7). The width of such extended nanostructures in each of the three cases was found to be between 8-10 nm (Figure 6.8). Aside from the extended nanostructures, the smaller unraveled structures also observed in all the three cases were < 50 nm in length and between 5 nm and 7 nm in width. Such small structures have been previously described for IF proteins and are thought to correspond to the protofibrils[13,14,44].

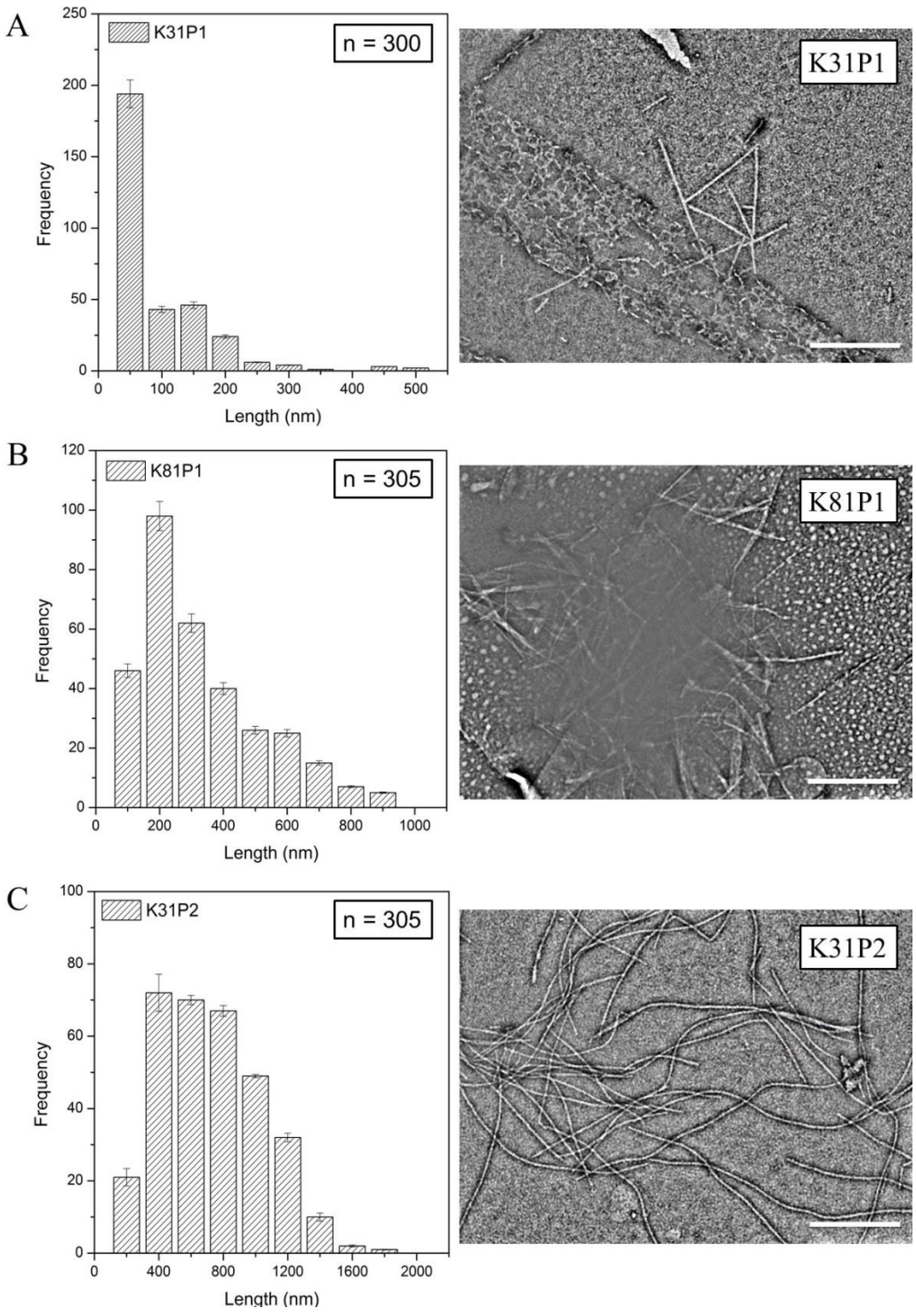


Figure 6.7: *Measurement of nanostructures of peptide acetates K31P1 (A), K81P1 (B) and K31P2 (C) at 10 mg/mL (w/v) in water stored at 25°C for over three weeks. Histograms used to present the length statistics have been obtained from nearly independent measurements. Panel shows the representative TEM micrographs. Error bars correspond to standard errors of the means of 10 independent measurements. (Scale bar is 200 nm)*

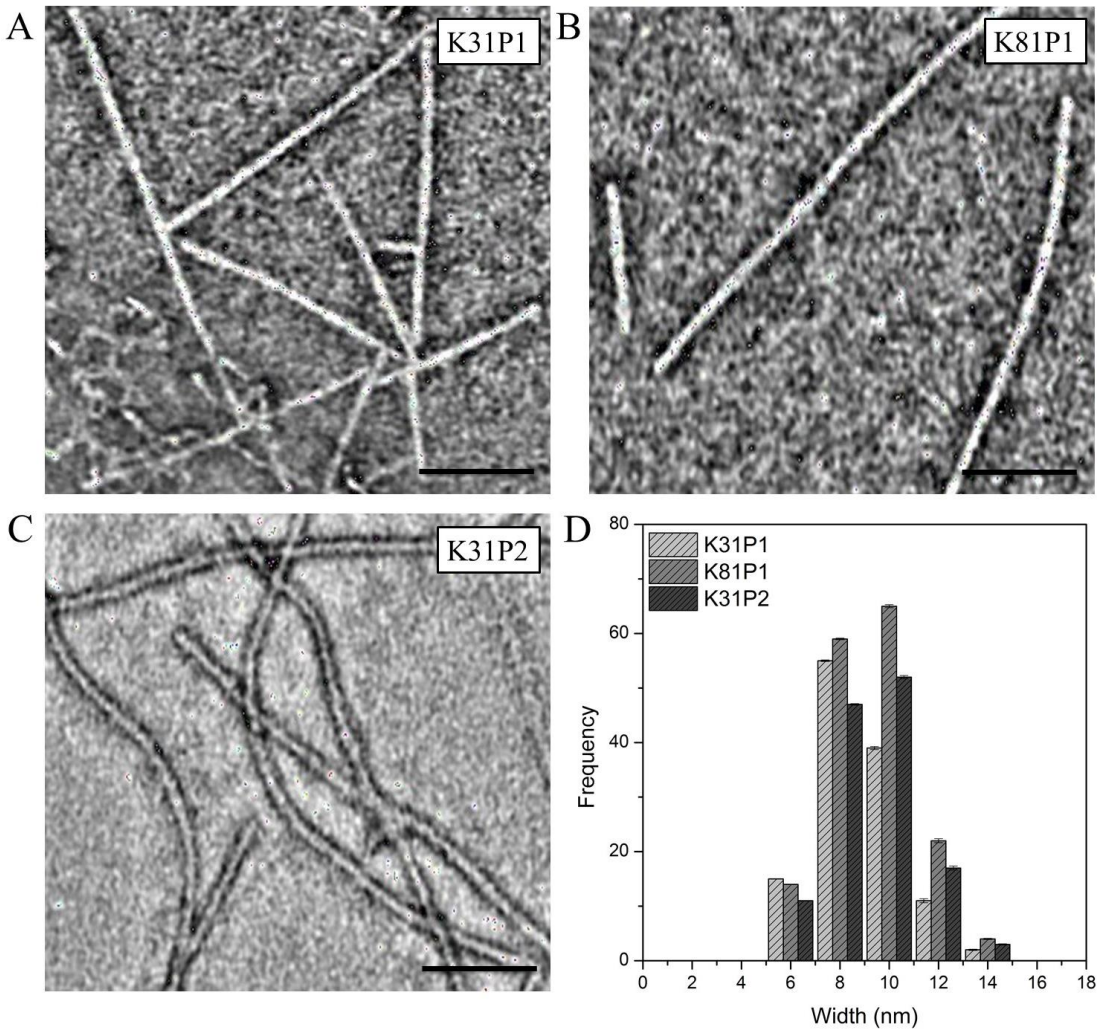


Figure 6.8: *Zoomed TEM micrographs of peptide acetates K31P1 (A), K81P1 (B) and K31P2 (C) at 10 mg/mL in water. (D) Histogram shows the width distribution of the extended filament like nanostructures collected for 300 particles. Red lines indicate the Error bars correspond to standard errors of the means of 10 independent measurements. (Scale bar is 50 nm)*

These results showed that the peptide sequences designed from the consensus sequence motif of type I and type II hard α-keratin specific to wool fibers self-assembled in isolation into nanostructures with extended morphology. Further, the self-assembly of peptides involved formation of short rod like structures which further self-assembled into extended filaments with dimensions similar to those described for intermediate filaments of full length keratins[14,15].

6.4.3 Effect of pH on peptide self-assembly

Solutions conditions like pH, ionic strength and the presence of reducing agents can affect the solubility of peptide acetates in solution[45,46]. In the case of mammalian epidermal keratins, basic pH and non-reducing conditions are known to favor filament assembly and thus decrease solubility[18,47]. The presence of a large number of charged residues distributed across the peptide sequences investigated make them susceptible to change in solution pH. Protonation and deprotonation as a function of pH is known to affect peptide the self-assembly process[48].

As mentioned above, the salt bridge between the glutamic acid residue of strand 1 and the arginine residue of strand 2 and the hydrogen bond between the tyrosine residue of strand 1 and the arginine residue of strand 2 within the consensus sequence motif are highly conserved interactions of the coiled-coil protein interface[32] (Figure 6.2 B). Such specific and essential non-covalent interactions of the parent interface can be conserved in the self-assembly process of the isolated peptide sequences. Therefore, the self-assembly of the peptides was studied:

(I) At different pH conditions (4.0 to 7.6) and in presence of salt (150 mM sodium chloride)

Peptide acetate samples prepared in different buffer solutions with pH in the 4.0 to 7.6 range and in buffer with 150 mM sodium chloride were monitored overtime. Peptide acetate K31P1 was found to soluble over the pH 4.0 to 7.6 range. Absence of any particulate matter or turbidity suggested no self-assembly of the peptide acetates (Figure 6.9 A).

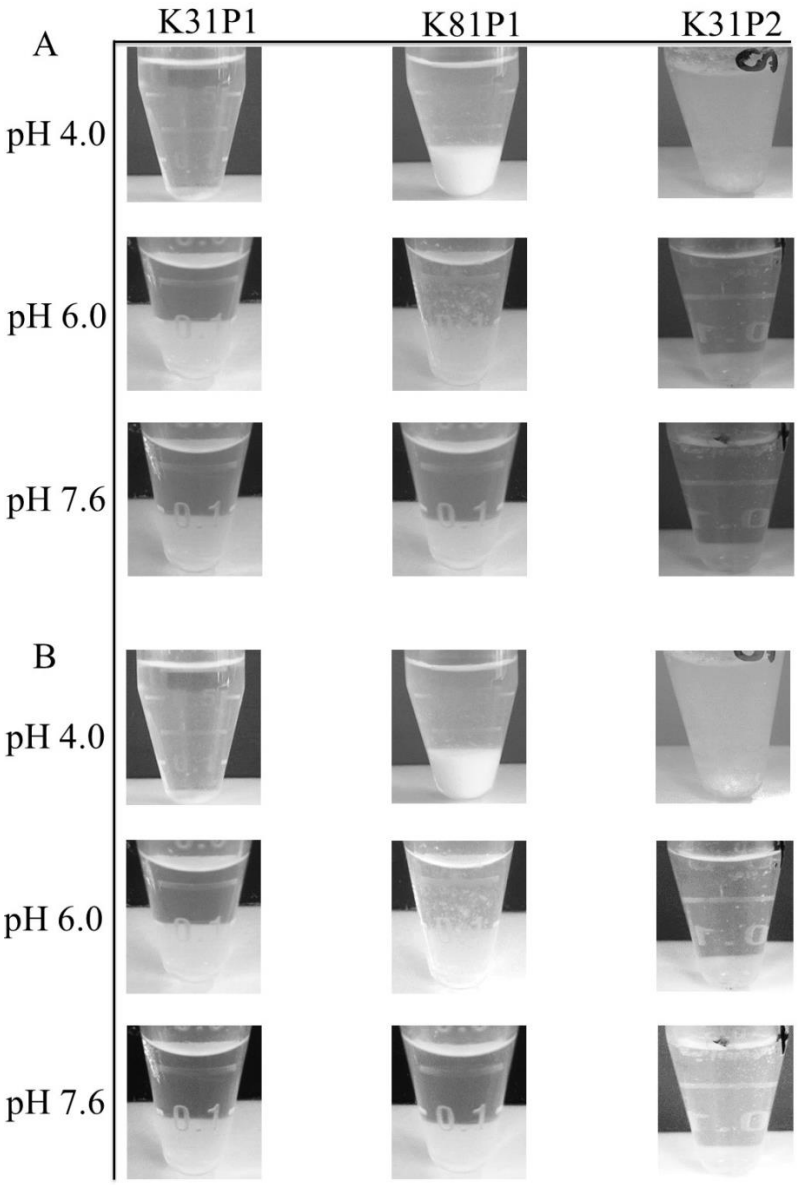


Figure 6.9: *Peptide acetates at 10 mg/mL at different pH conditions (A) and in buffer containing 150 mM sodium chloride (B) stored for 3 weeks at 25°C. pH 4.0: 20 mM sodium acetate; pH 5.0: 20 mM sodium acetate; pH 6.0: 20 mM MES; pH 7.6 20 mM Tris-HCl.*

Peptide acetates K81P1 and K31P2 were soluble only under basic pH, however, at acidic pH < 6.0 visible precipitates observed for both K81P1 and K31P2 peptides indicated a saturated peptide solution. This decrease in solubility of the peptide acetates can be explained by the solution pH nearing the iso-electric point (pI) of the peptide acetates, which falls between 3.5 and 4.6 for all the three peptides.

The turbid appearance of the peptide acetate samples, similar to that observed for peptide acetates in water (Figure 6.9 A) however, suggested the formation of extended peptide assemblies. Peptide acetate samples prepared in buffer containing 150 mM sodium chloride were similar in appearance to the buffer only samples and did not show any substantial difference in solubility (Figure 6.9 B).

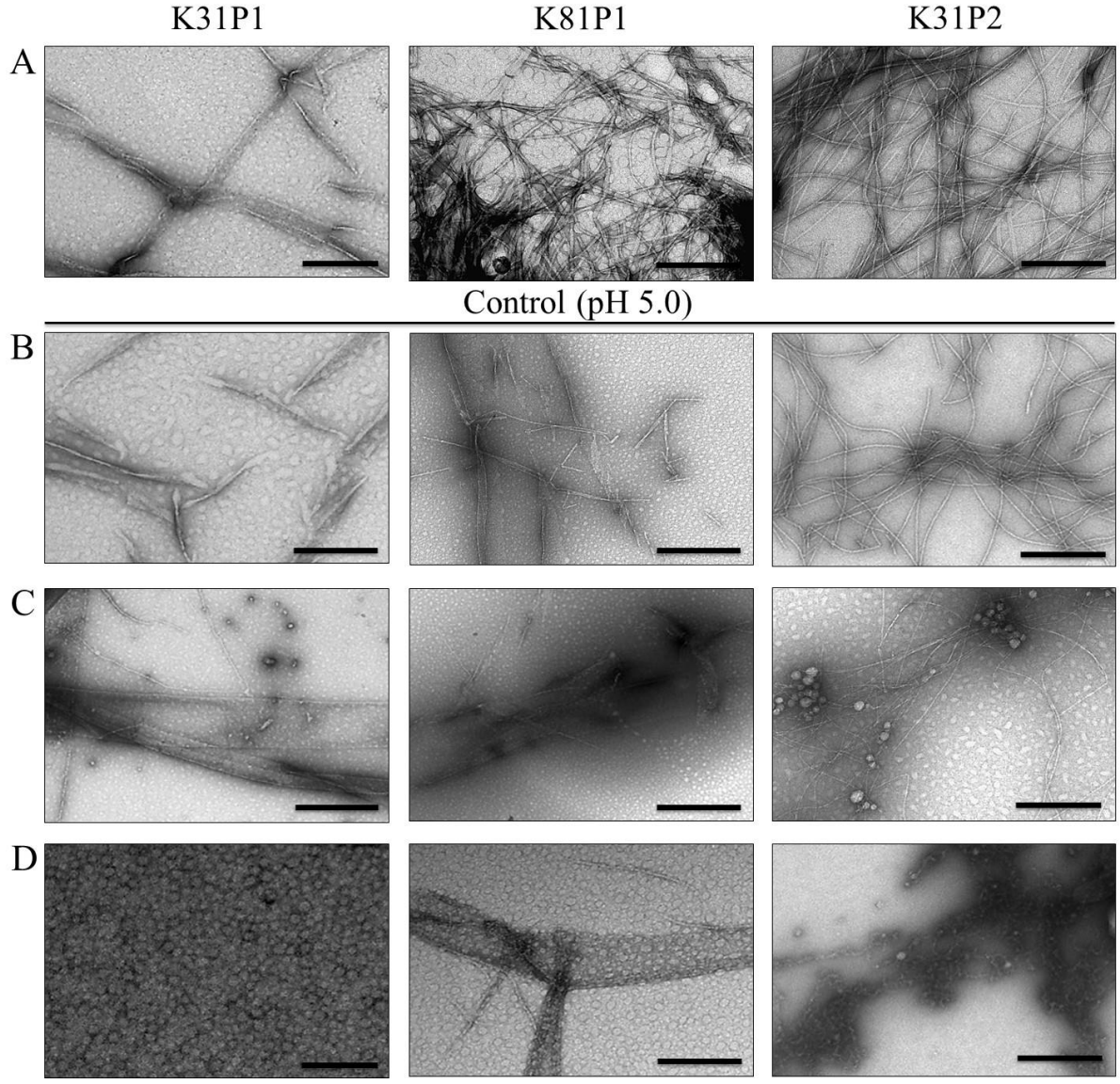


Figure 6.10: *Negative stained transmission electron micrographs of peptide acetates at 10 mg/mL in 20 mM buffer solution at different pHs stored at 25°C in over three weeks. (A) 20 mM sodium acetate pH 4.0 ;(B) 20 mM sodium acetate pH 5.0 ;(C) 20 mM MES pH 6.0 ;(D) 20 mM Tris-HCl pH 7.6. (Scale bar is 200 nm)*

Electron microscopy analysis of the peptide acetate samples in buffer at near neutral pH of 7.6 revealed that the K31P2 peptide self-assembled into nanostructures, with the presence of short rod like fragments (Figure 6.10 D). Such assemblies with extended morphology have been previously described for synthetic peptide designed from consensus sequence motif mixed with epidermal keratins[41]. No such assembly was observed for either of the shorter peptides K31P1 and K81P1 under these conditions.

At a slightly acidic pH of 6.0, all three peptides self-assembled in solution with short rod like fragments and wide extended tapes observed for both K31P1 and K81P1, and extended filament like nanostructures observed for the K31P2 peptide (Figure 6.10 C). The extended filament assembly of the K31P2 peptide was found to be stable over the pH range 4.0 to 6.0 (Figure 6.10 A-C). In the case of the K81P1 peptide, the extended filament like assembly was favored under more acidic pH 5.0, consistent with that observed for the peptide acetate in water sample (Section 6.3.2). At pH 4.0, the specific self-assembly of the K81P1 peptide was found to be more pronounced with the presence of extended filament like nanostructures (Figure 6.10 A). For the shorter peptide K31P1, no specific extended filament like assembly was observed under the conditions investigated. However, the presence of short rod like fragments and nanostructures with tape like morphology, consistent at all acidic pH conditions trialed, suggested that the K31P1 peptide self-assembled in solution.

Measurements performed on the nanostructures observed for the K31P2 peptide at pH 4.0 were on an average ~ 800 to 1000 nm in length and ~ 10 nm in width, consistent with those observed for control pH sample (Figure 6.11 C).

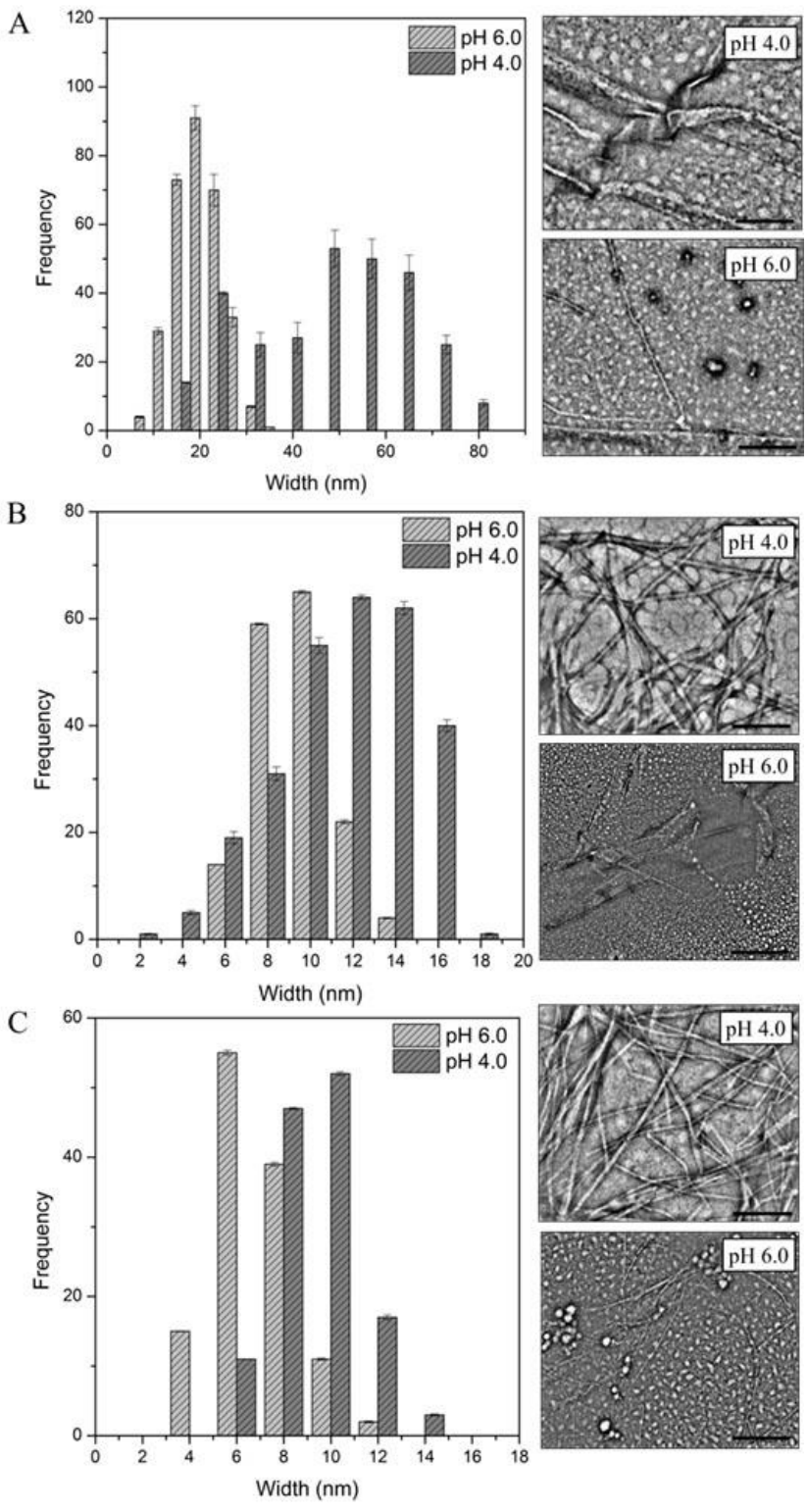


Figure 6.11: *Measurement of nanostructures of peptide acetates K31P1 (A), K81P1 (B) and K31P2 (C) at 10 mg/mL in buffer at different pH. Panel shows the representative TEM micrographs. Error bars correspond to standard errors of the means of 10 independent measurements. (Scale bar is 50 nm).*

The histogram shows the distribution of the width the different self-assembled peptide nanostructures collected from ~ 300 particle measurements done using ImageJ software. The zoomed EM micrograph is produced using ImageJ by applying bandfass filter, shows the details of the peptide nanostructures.

However, at higher pH 6.0 the extended nanostructures were less wide ~ 5 nm, which suggested that the specific self-assembly of the K31P2 peptide into extended filaments similar in dimension to keratin intermediate filaments is favored at low pH $\leq$ 5.0[18,49]. In the case of K81P1, the extended filament assemblies observed at pH 4.0 were similar in dimension to those observed for the control pH sample, consistent with the K31P2 peptide. Measurements performed on the nanostructures observed for the shorter K31P1 peptide at pH 6.0 revealed that the peptide nanotapes were 60-90 nm in length.

6.4.4 Effect of redox state on peptide self-assembly

As mentioned earlier, disulfide crosslinking of keratin oligomers is inherent to the self-assembly of hard α-keratins into cross-linked fibrous networks[15,32]. However, the significance of disulfide crosslinking between cysteine residues that are part of the coiled coil protein-protein interface at the lower level of self-assembly is not well understood. The peptide sequences derived from the consensus sequence motif of wool hard α-keratins type I and type II, were modified by addition of a cysteine residue at both ends of the peptide. This was done with an aim to build peptide tectons that can be triggered for self-assembly. Therefore, to study the effect of change in peptide assembly as a function of redox state of the peptide, the experiments were conducted with peptide acetates dissolved in buffer containing TCEP (reducing conditions) and buffer containing hydrogen peroxide (oxidising conditions).

Earlier results showed that all the three peptides under non-reducing conditions and in the presence of air, and under acidic conditions self-assembled into nanostructures (Section 6.3.2). As the peptide assembly was found to be sensitive to pH change, the experiments were carried out in the solution most favorable for peptide self-assembly (pH 5.0 sodium acetate buffer at 25℃). The control peptide sample for each of the three peptides was prepared by dissolving the lyophilized peptide acetate in pH 5.0 buffer devoid of any oxidising or reducing agent.

Peptide acetate samples stored at 25℃ for over three weeks for each of the conditions investigated self-assembled into nanostructures, with extended tape like assembly observed for K31P1 peptide and extended filaments observed for both K81P1 and K31P2 peptides (Figure 6.12). The extent of self-assembly of the peptides into tapes or extended filaments for each of the three peptides, however was found to be favored in presence of a 10-fold molar excess of hydrogen peroxide (oxidising conditions). For peptide acetates dissolved in buffer containing TCEP, no specific tape like assembly was observed for the K31P1 peptide, though the appearance was consistent with the presence of relatively smaller nanostructures. Both K81P1 and K31P2 peptides self-assembled into extended filament like nanostructures, however these filaments were larger in width and appear unraveled than the extended filaments observed under oxidising conditions for the peptides (Figure 6.12, B and C). Keratin filaments under TEM are usually characterized as smooth, unbranched, flexible individual filaments that may be sometimes bundled. Unraveled filaments usually correspond to disentangled protofibrils that are no longer in tight contact with each other along the length of the filament and thus appear wider. Furthermore, long stretches of unraveled filaments are interspersed with shorter more compact stretches[50].

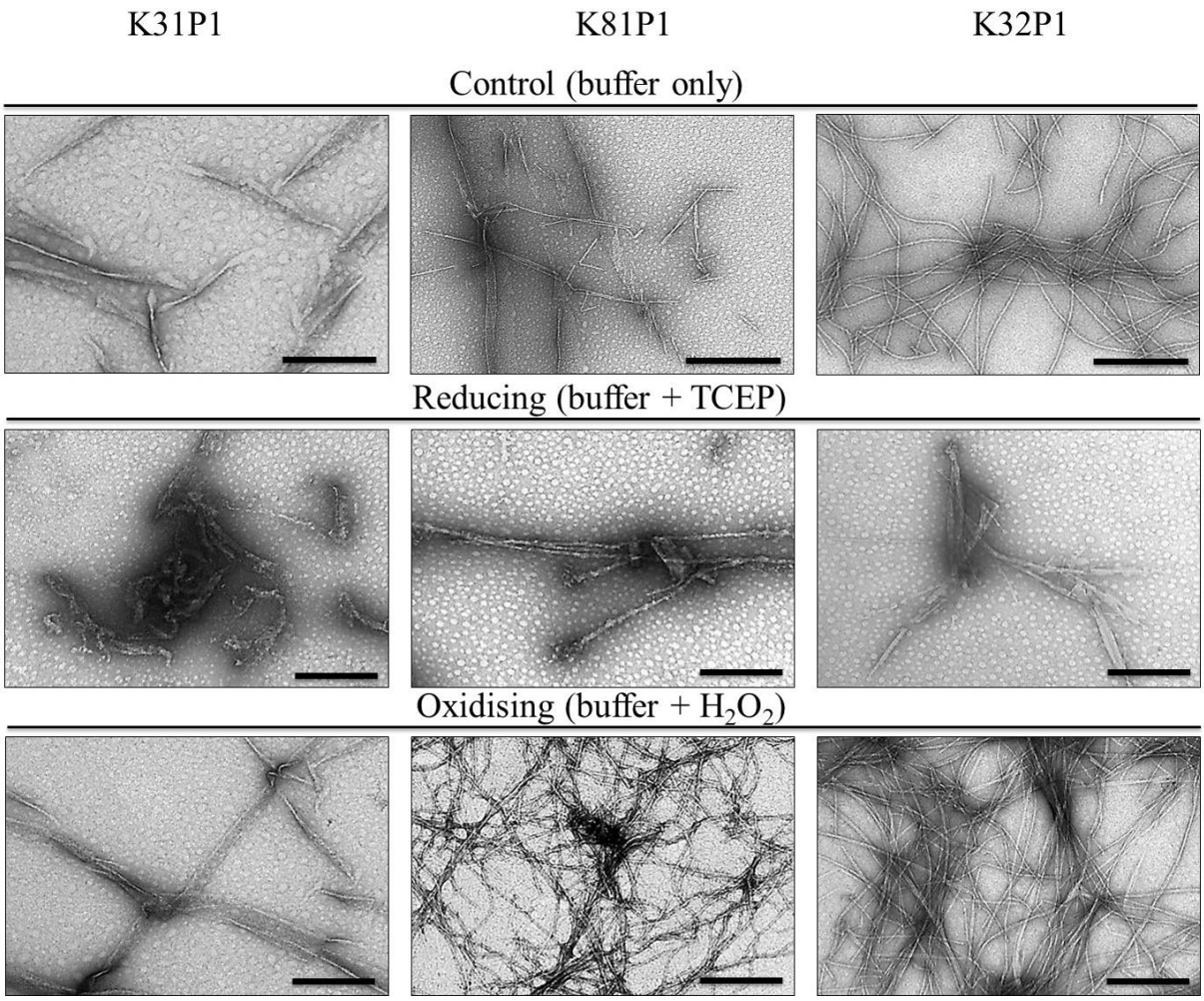


Figure 6.12: *TEM of peptide acetates in reducing and oxidising conditions: Negative stained electron micrographs of peptide acetates K31P1 , K81P1 and K31P2 dissolved in pH 5.0 buffer (buffer only);peptide acetate dissolved in pH 5.0 buffer with TCEP (Reducing) and peptide acetate in pH 5.0 with H_2O_2 (oxidising). TEM micrographs for peptide acetate samples stored at 25°C were taken at the end of three weeks. (Scale bar is 200 nm)*

To further examine the effect of reducing conditions (TCEP) on peptide self-assembly, the self-assembled extended peptide nanostructures formed under oxidising conditions were dialyzed into buffer containing TCEP. TEM analysis of such peptide samples revealed nanostructures with extended morphology for two of the three peptides (Figure 6.13).

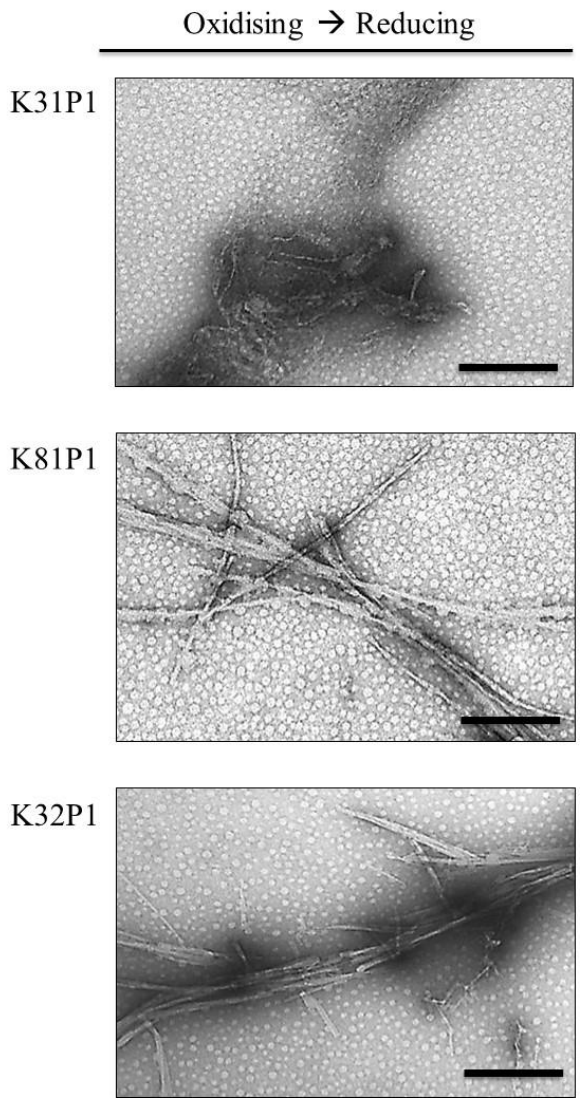


Figure 6.13: *TEM of peptide acetates under reducing conditions: Negative stained electron micrographs of peptide acetates K31P1 , K81P1 and K31P2 dissolved in pH 5.0 with H_2O_2 (oxidising) stored over three weeks at 25°C dialyzed into pH 5.0 buffer with TCEP. TEM micrographs for peptide acetate were taken 24 hours after dialysis into reducing buffer. Both K81P1 (B) peptide and K31P2 peptide (C) form extended nanostructures under these conditions. (Scale bar is 200 nm)*

Measurement of the peptide nanostructures, revealed that the structures observed for K31P1 peptide were smaller in size and much wider in dimension than the peptide nanotapes observed under oxidising conditions (Figure 6.14).

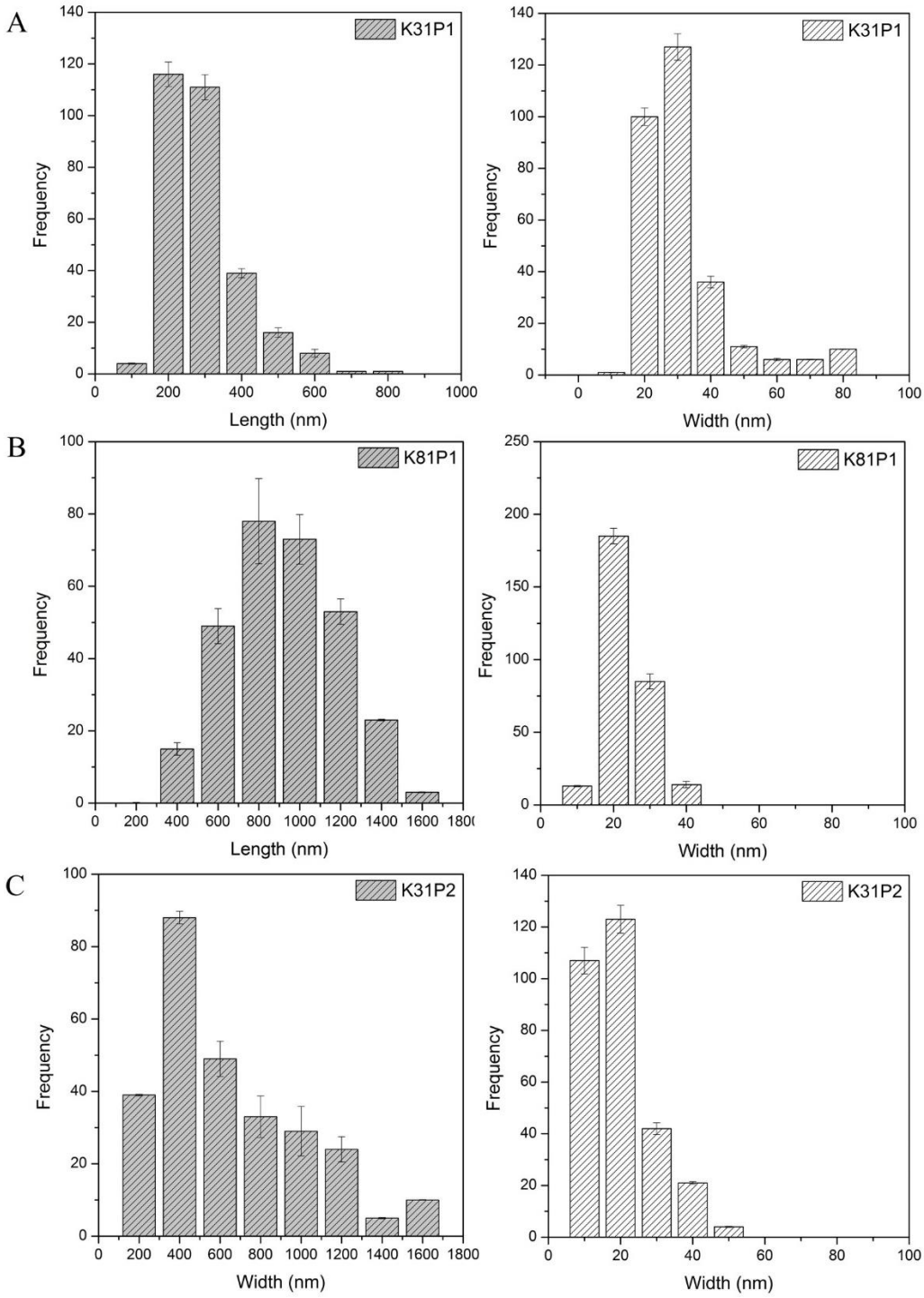


Figure 6.14: *Measurement of nanostructures formed by peptide acetates K31P1 (A), K81P1 (B) and K31P2 (C) at 10 mg/mL dissolved in buffer with TCEP pH 5.0 as shown in figure 6.13. 300 measurements from the data set for the peptide at the conditions investigated were used to present length and width statistics. Error bars correspond to standard errors of the means of 10 independent measurements.*

In the case of the K81P1 peptide the extended nanostructures were similar in length but nearly twice as wide. For the K31P2 peptide the nanostructures were much smaller in length compared to those formed under non-reducing conditions and in presence of hydrogen peroxide. These results suggested that disulfide cross-linking is not necessary for peptide self-assembly. The differences in the overall morphologies and dimensions of the peptide assemblies observed under non-reducing conditions and under oxidising conditions, however, suggest a role for cysteine residues in the specific self-assembly of the peptides into extended tapes or individual extended filaments.

The absence of the specific self-assembly of K31P1 peptide into upon dialyzing into buffer with TCEP suggested that the peptide self-assembly into nanotapes is not favored under reducing conditions. These peptide nanostructures were similar in dimension to those observed under reducing conditions. TEM of K81P1 and K31P2 peptide acetates dialyzed into reducing buffer revealed disordered peptide aggregates and some bundles of extended filaments. Such, bundles of filaments have been described for full length keratins subjected to proteolysis[14]. This indicated that the change in redox state of the peptide affects hierarchical self-assembly of the peptide.

To examine the morphology and size of the extended filaments formed by peptides K81P1 and K31P2 at acidic pH, under oxidising conditions, single particle analysis was performed from EM micrographs taken from different sample preparations (Figure 6.15).

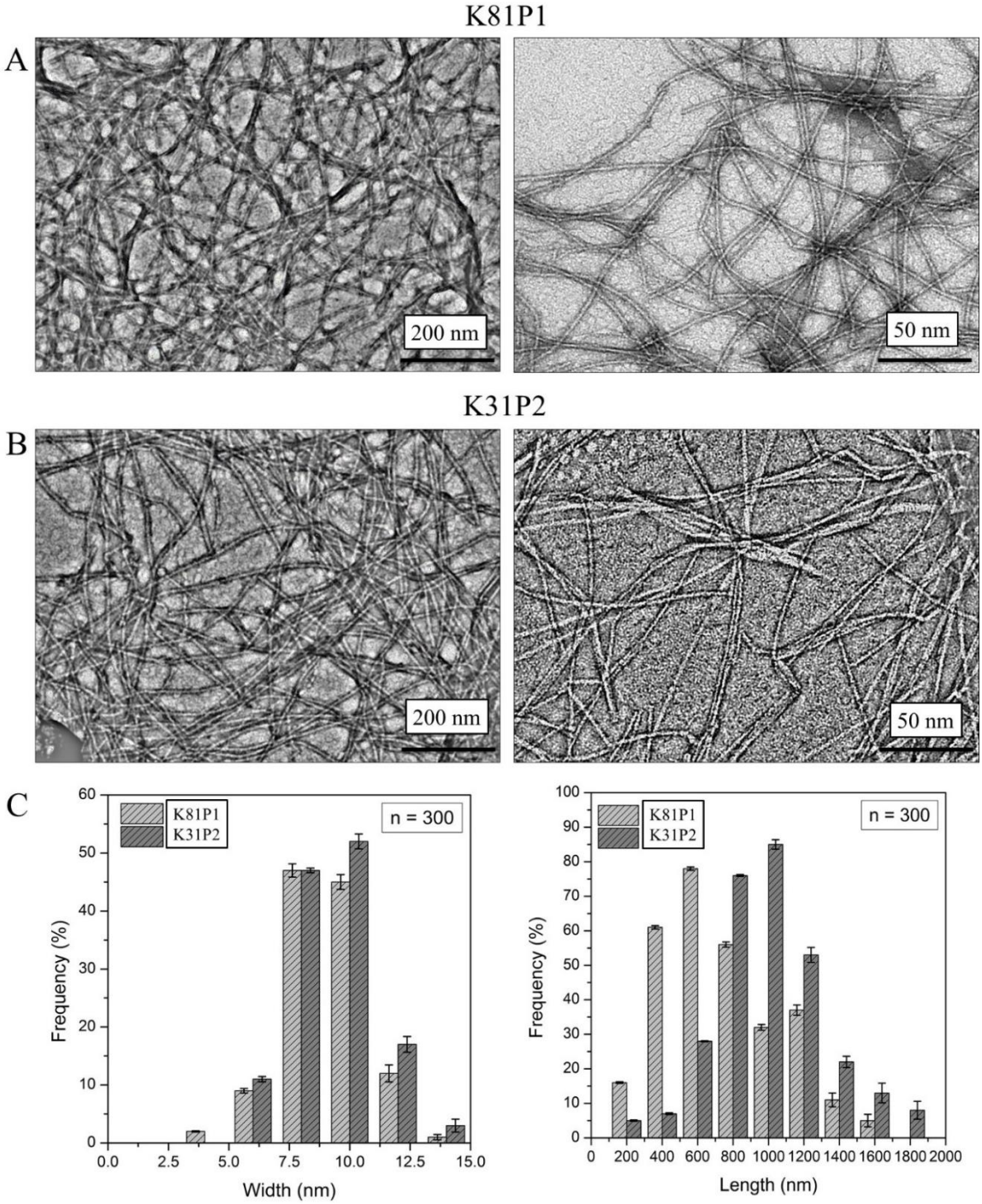


Figure 6.15: *Measurement of filament like nanostructures of peptide acetates K81P1 (A) and K31P2 (B) at 10 mg/mL in pH 5.0 buffer with H_2O_2 stored at 25°C for over three weeks. Data for measurements shown has been obtained from three different independent experiments. 300 measurements from the data set for each of the peptide acetates was used to present the width and length statistics. Very short nanostructures are <150 nm, short IFs are 200 - 500 nm, intermediate nanostructure are 500 to 1000 nm long and long nanostructure are > 1000 nm*[51].

Image analysis of the nanostructures done using ImageJ software revealed the average width of the nanostructures to be between 8 nm to 10 nm, with most of the filaments having a diameter of 10 nm. The size of the extended filaments under oxidising conditions was found to be larger than those observed for the peptide in water or pH 5.0 buffer with measured average overall length of ~600 nm for K81P1 peptide and ~1 µm for K31P2 peptide.

6.4.5 SDS-PAGE analysis of peptide nanostructures

To study the redox sensitivity of the peptides in solution, peptide samples prepared in pH 5.0 buffer stored at room temperature over 24 hours were analyzed by gel electrophoresis. Peptide solutions were centrifuged at 10,000 rpm for 10 minutes to pellet down suspended high molecular weight (HMW) peptide aggregates. The samples were run on gel (4% to 12%) containing sodium dodecyl sulfate (SDS) under reducing with TCEP and non-reducing without TCEP.

The peptide acetates were equilibrated at room temperature (~25°C) for 24 hours. To examine the effect of reducing conditions on peptide self-assembly, peptide acetates dissolved and equilibrated for 24 hours in pH 5.0 buffer were dialyzed into pH 5.0 buffer with TCEP overnight. Peptide acetate samples were mixed with PAGE loading buffer with TCEP, while other samples were mixed with non-reducing PAGE loading buffer in a 1:1 ratio. Figure 6.16 shows the comparison of peptide acetates under the different conditions.

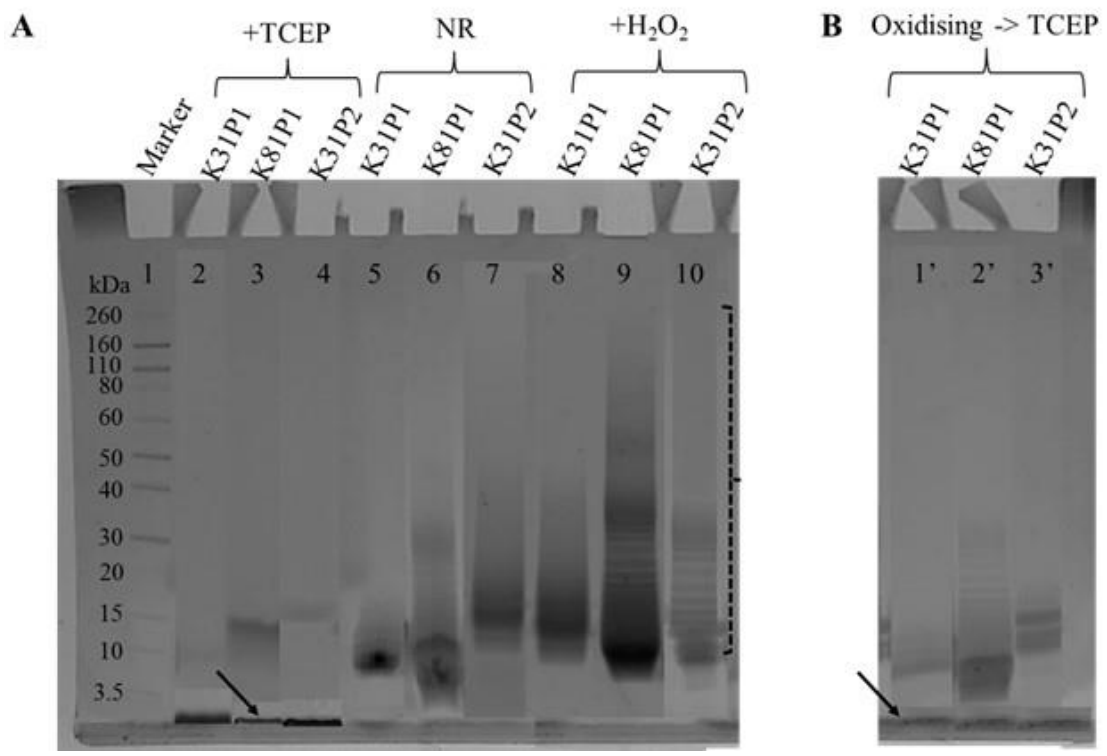


Figure 6.16: *SDS-PAGE of the peptide acetates in pH 5.0 buffer under different conditions: +TCEP, peptide acetate in buffer with 10 mM TCEP (reducing conditions); NR, peptide in buffer only (non-reducing conditions); + H_2O_2, peptide acetate in buffer containing 100 mM of hydrogen peroxide (oxidising condition); oxidising -> TCEP, peptide acetate sample equilibrated under oxidising conditions dialyzed into reducing buffer. Black arrows indicate the low molecular weight peptide monomer of ~ 1.4 kDa and dashed blue bracket indicate the faint ladder like bands for peptide acetate equilibrated under oxidising conditions.*

Note: The SDS-PAGE gel shown in Figure 6.16 A has been compiled from three different gel analysis to show the behavior of the peptide acetates in gel for the different conditions investigated.

The reducing gel (Lane 2, 3, 4 in Figure 6.16) showed an intense band near the 15 kDa molecular marker and faint bands at low molecular weights which suggested that these band corresponded to peptide that is not cross-linked by disulfide bonds (possible SDS-resistant species). Further, the presence of higher intensity band below 3.5 kDa molecular weight for all

the three peptides under reducing conditions suggested that the most of the peptide in solution existed in the monomer form (M_w of K31P1: 1.3 kDa; K81P1: 1.4 kDa; K31P2: 1.7 kDa).

In the non-reducing gel (lanes 5, 6 and 7 in Figure 6.16) a number of faint bands observed between 20 kDa to 3 kDa for peptide K31P1; between 50 kDa to 3.5 kDa for peptide K81P1 and peptide K31P2 suggested the presence of peptide oligomers. Additionally, peptide acetate samples equilibrated in buffer with 100 mM hydrogen peroxide, showed several faint bands with molecular weights above the 15 kDa band. Further, low molecular weight bands corresponding to smaller peptide oligomer appeared faint under non-reducing conditions, but gain considerable intensity under reducing conditions as evident in lane 2,3 and 4 (Figure 6.16). These results suggested that under non-reducing conditions (pH 5.0 buffer only) and oxidising condition (pH 5.0 buffer + 100 mM hydrogen peroxide) the peptide oligomer contains disulfide bonds. The current results show that TCEP was effective in breaking the disulfide linkage resulting in LMW bands compared to the non-reducing and oxidized sample for each of the peptide acetates.

Under oxidising conditions all three peptide sequences exhibit covalently linked oligomers visualized as multiple bands in the gel (lane 8, lane 9 and lane 10). As the gel and sample loading dye contains the detergent SDS, most of the non-covalent interactions between the peptide molecules are disrupted[51]. All the three peptide sequences contain a cysteine residue present at both the N- and C- terminal end of the peptide sequence with the respective amino group and carboxyl group capped. The absence of reducing agent and/or presence of oxidising agent favors the formation of inter-peptide disulfide bonds between cysteine residues.

Additionally, the LMW bands observed for the peptide samples under non-reducing conditions are faint (Lane 5,6 and 7) and develop considerably in intensity under reducing conditions (lane 2,3 and 4). Bands corresponding to HMW oligomers observed for peptide K31P1, peptide

K81P1 and K31P2 (lane 5, lane 6 and lane 7) under non-reducing conditions further suggested that these bands could correspond to disulfide linked peptide oligomers.

The faint HMW bands observed for the peptide samples under reducing conditions (lane 2,3 and 4) consistent with the peptide samples under non-reducing conditions can be explained by the formation of SDS-resistant species. SDS resistant complexes can result from formation of strong hydrophobic interactions or non-disulfide cross-links. Studies focused on such protein complexes identified in context of important biological functions, have highlighted the role of some critical amino acid residues and the molecular interactions involved[52,53]. Mass spectrometry analysis of peptide acetates equilibrated under reducing and oxidising conditions would therefore need to done to confirm the identity of the bands observed in SDS-PAGE gel under both reducing and oxidising conditions.

Overall the SDS-PAGE results are consistent with the extended nanostructure assembly observed under oxidising condition by TEM. The HMW aggregates (M_r over 260 kDa) and LMW aggregates (M_r between 10 kDa and 40 kDa) observed as the ladder for all the three peptides potentially suggest a hierarchical extended filament like nanostructure with oligomeric units that are resistant to SDS.

6.5 Discussion

Peptide sequences with a heptad pattern show a natural propensity to exist in a helical conformation in solution and can self-assemble into a range of extended fibrous nanostructures[54]. In this work, peptide sequences (K31P1, K81P1 and K31P2) corresponding to the consensus sequence motif of *ovine* type I and type II hard α-keratins flanked with cysteine at both N- and C-terminus were investigated as molecular tectons.

All the three peptide sequences showed propensity to exist in the parent conformation as a α-helix in solution indicated by the double minima at 208 nm and 222 nm. Peptide acetates dissolved in water to assess the self-assembly of the peptide in solution showed visible precipitates for two out of the three peptides: K81P1 and K31P2. Electron microscopy analysis of these peptide acetate samples revealed that all the three peptides stored at room temperature self-assembled over time with mostly short rod shaped structures for the K31P1 peptide and K81P1 peptide, and bundles of extended filament like nanostructures for K31P2 peptide. These conditions are prone to allow inter- and intra-molecular disulfide crosslinking over time as reported by Sun and Green in 1978 for extracted epidermal keratins in filament assembly in the presence of air[55].

Protonation and deprotonation of charged residues as a function of pH is known to influence the conformation of peptides in solution and subsequent self-assembly. In the case of keratins pH of 7.6 is known to favor the self-assembly into filaments. Samples of peptide acetates dissolved in Tris-HCl buffer at pH 7.6 were found to be soluble in each of the three cases with no visible particulates. EM revealed that only peptide K31P2 self-assembled, with mostly short rod like fragments and thin filaments predominant under this condition. Peptide acetates dissolved in acidic buffer, however, were found to be less soluble with visible precipitates observed for peptides K81P1 and K31P2 below pH 6.0. Acidic conditions favored the self-assembly of the peptides into extended tape like filaments of the K31P1 peptide at and below pH 5.0; bundles of extended filament like nanostructures for both peptide K81P1 and K31P2 at pH 4.0 and pH 5.0 respectively. Assembly of IF proteins including keratins is reported to occur at near neutral or alkaline pH and ionic strength between 10 mM to 50 mM[16].

The self-assembly of the peptide sequences with their design inspired from the IF consensus sequence motif into extended filaments favored under acidic condition has previously not been observed. The C-terminus coil 2B region of the central coiled coil rod domain of IF proteins including keratins shows more negative surface potential, arising from the localization of glutamic acid and aspartic residues in this region. Therefore, neutral or slightly alkaline pH would favor the non-covalent interactions between the two protein helical domains. However, acidic pH seems to be an important driving force in the self-assembly of these peptide sequences as no significant difference is observed for filament like nanostructures for peptides in water and in 25 mM ionic strength buffer.

Further, the self-assembly of the isolated peptide sequences under more acidic conditions suggested that the complementary non-covalent interactions of the parent interface were retained during peptide self-assembly. Also, the protonation of the charged residues: aspartic acid (E), glutamic acid (D), asparagine (R) and lysine (K) as a function of pH did not substantially affect the peptide self-assembly process.

Studies on native epidermal keratin and recombinant keratin with an incorporated cysteine residue showed that in absence of reducing agent, and upon exposure to air, disulfide bonded HMW keratin aggregates form[32]. Also, when such a keratin sample was left for five days under the same conditions, very high HMW aggregates formed that could not enter the gel[55]. When such aggregates are exposed a reducing environment in the presence of DTT or β-mercaptoethanol, LMW protein aggregates and monomers are present confirming disulfide linked keratin oligomers[55].

Study of the designed peptide sequences under oxidising and reducing conditions was suggestive of disulfide linking between the peptides. Samples of the peptide in presence of TCEP did not favor the specific self-assembly of the peptide under acidic conditions. However, in the absence of the reducing agent and on exposure to air, all the three peptides self-assembled in solution with extended filament like nanostructures observed for two out of the three peptides (K81P1 and K31P2).

SDS-PAGE gel electrophoresis of the peptide acetate samples stored in buffer with excess of hydrogen peroxide revealed a smear of multiple bands that collapse upon reducing with TCEP, further supported the presence of disulfide linking in the peptide assembly. Hydrogen peroxide added to the peptide solution at ~10 fold molar excess to the peptide resulted in the self-assembly of the peptide in solution into extended filaments. The ladder like bands observed, specifically for peptides K81P1 and K31P2 under oxidising conditions bear similarity to the multiple bands observed for cross-linked fibrinogen and bis(Gly-Pro-Arg-Pro)peg complexes[56] and heated whey protein isolates[57]. Collapse of these ladder like bands upon subjecting to reducing conditions indicated disulphide linking. Similar behavior was observed for the peptide sequences under non-reducing and reducing SDS-PAGE, which supports the inference that within the extended nanostructures the peptide is disulfide cross-linked.

The results showed that the terminal cysteine residues take part in chemical cross-linking of the peptide into oligomers and in the elongation of the peptide oligomers into short rod like filaments and extended filaments. Further, the smear of closely placed ladder like multiple bands observed on SDS-PAGE gel under non-reducing conditions and oxidsing conditions for the peptides suggested a linear daisy chain like. This type of crosslinking could potentially arise from the end to end association of peptide oligomers.

These findings presents a potential alternate mode of molecular self- assembly of keratin derived peptide motifs *in vitro,* mediated by cysteine residues. This is different to the earlier reported work by Steinert *et al*., 1976 and a recent discovery of a transdimer disulfide bond *in vitro* occurring during keratin assembly which limits filament elongation and impairs filament maturity [32,50]. Mature filaments assembled from heteromeric assembly of K8-K18 and K5-K14 are reported to be a smooth extended rod shape [41,58]. The extended filament like assemblies observed for the peptides investigated in this work appear to be smooth, similar to the reported IFs formed by the *in vitro* assembly of full length epidermal keratins. The results of this work show that peptide sequences of the IF consensus motif inspired from cuticular keratins can interact in a complementary fashion to each other as homo-oligomers and further self-assemble into extended filament like nanostructures. Such filament nanostructures are comparable to the mature wild type filaments formed by type I and type II keratins assembling as heterodimers.

These findings indicate that the highly conserved consensus sequence motif of keratins shows a propensity to assemble as a homomer, not reported previously for keratin protein assembly. The results allow for some speculation for keratin homodimer assembly to be present *in vitro* and even in vivo. As the type I and type II keratins are differentially expressed in a differentiating cuticle. The cysteine residues found naturally flanking the consensus motif at the C-terminus end of the coil 2 B motif of the central helical rod domain found specifically in type I can potentially be involved in disulfide linking.

To build further on our findings in the future, a detailed study of the peptide sequences by mass spectrometry, molecular dynamic simulation and crystallization, if possible, would allow for a more realistic view of the homomer assembly of the peptide sequences and the predicted intramolecular disulfide bonds.

6.6 References

1. Woolfson, D. N., Bartlett, G. J., Bruning, M. & Thomson, A. R. (2012). New currency for old rope: from coiled-coil assemblies to α-helical barrels. *Current Opinion in Structural Biology* **22**, 432-441.
2. Rackham, O. J., Madera, M., Armstrong, C. T., Vincent, T. L., Woolfson, D. N. & Gough, J. (2010). The evolution and structure prediction of coiled coils across all genomes. *Journal of Molecular Biology* **403**, 480-493.
3. Gruber, M. & Lupas, A. N. (2003). Historical review: another 50th anniversary--new periodicities in coiled coils. *Trends in Biochemical Science* **28**, 679-685.
4. Parry, D. A., Fraser, R. D. & Squire, J. M. (2008). Fifty years of coiled-coils and α-helical bundles: a close relationship between sequence and structure. *Journal of Structural Biology* **163**, 258-269.
5. Crick, F. (1953). The packing of [α]-helices: simple coiled-coils. *Acta Crystallographica* **6**, 689-697.
6. Mason, J. M. & Arndt, K. M. (2004). Coiled coil domains: stability, specificity, and biological implications. *Chembiochem* **5**, 170-176.
7. Wang, X., Kim, H. J., Wong, C., Vepari, C., Matsumoto, A. & Kaplan, D. L. (2006). Fibrous proteins and tissue engineering. *Materials Today* **9**, 44-53.
8. Yigit, S., Dinjaski, N. & Kaplan, D. L. (2016). Fibrous proteins: at the crossroads of genetic engineering and biotechnological applications. *Biotechnology and Bioengineering* **113**, 913-929.
9. Feughelman, M. (1997). Mechanical properties and structure of α-keratin fibres: wool, human hair and related fibres.*UNSW Press*.
10. Vasconcelos, A., Freddi, G. & Cavaco-Paulo, A. (2008). Biodegradable materials based on silk fibroin and keratin. *Biomacromolecules* **9**, 1299-1305.
11. Lee, H., Noh, K., Lee, S. C., Kwon, I.-K., Han, D.-W., Lee, I.-S. & Hwang, Y.-S. (2014). Human hair keratin and its-based biomaterials for biomedical applications. *Tissue Engineering and Regenerative Medicine* **11**, 255-265.
12. Mogosanu, G. D., Grumezescu, A. M. & Chifiriuc, M. C. (2014). Keratin-based biomaterials for biomedical applications. *Current Drug Targets* **15**, 518-530.
13. Aebi, U., Häner, M., Troncoso, J., Eichner, R. & Engel, A. (1988). Unifying principles in intermediate filament (IF) structure and assembly. *Protoplasma* **145**, 73-81.

14. Aebi, U., Fowler, W. E., Rew, P. & Sun, T. T. (1983). The fibrillar substructure of keratin filaments unraveled. *Journal of Cell Biology* **97**, 1131-1143.
15. Herrmann, H., Strelkov, S. V., Burkhard, P. & Aebi, U. (2009). Intermediate filaments: primary determinants of cell architecture and plasticity. *The Journal of Clinical Investigation* **119**, 1772-1783.
16. Herrmann, H. & Aebi, U. (2004). Intermediate filaments: molecular structure, assembly mechanism, and integration into functionally distinct intracellular Scaffolds. *Annual Review in Biochemistry* **73**, 749-789.
17. Steinert, P. M., Wantz, M. L. & Idler, W. W. (1982). O-phosphoserine content of intermediate filament subunits. *Biochemistry* **21**, 177-183.
18. Steinert, P. M., Idler, W. W. & Zimmerman, S. B. (1976). Self-assembly of bovine epidermal keratin filaments *in vitro*. *Journal of Molecular Biology* **108**, 547-567.
19. Steinert, P. M. (2001). Keratins: Dynamic, flexible structural proteins of epithelial cells. *Current Problems in Dermatology* **13**, 193-198.
20. Coulombe, P. A. & Omary, M. B. (2002). 'Hard' and 'soft' principles defining the structure, function and regulation of keratin intermediate filaments. *Current Opinion in Cell Biology* **14**, 110-122.
21. Bragulla, H. H. & Homberger, D. G. (2009). Structure and functions of keratin proteins in simple, stratified, keratinized and cornified epithelia. *Journal of Anatomy* **214**, 516-559.
22. R.D. Fraser, H. Z. R. a. B. L. (2015). The structure of a marino wool fibre. *CSIRO*.
23. Herrmann, H., Häner, M., Brettel, M., Ku, N.-O. & Aebi, U. (1999). Characterization of distinct early assembly units of different intermediate filament proteins1. *Journal of Molecular Biology* **286**, 1403-1420.
24. Herrmann, H., Wedig, T., Porter, R. M., Lane, E. B. & Aebi, U. (2002). Characterization of early assembly intermediates of recombinant human keratins. *Journal of Structural Biology* **137**, 82-96.
25. BeefLambNZ. (2014). Compendium of New Zealand Farm Facts. *BeefLambNZ*.
26. Brown, J. H., Cohen, C. & Parry, D. A. (1996). Heptad breaks in α-helical coiled coils: stutters and stammers. *Proteins* **26**, 134-145.
27. Herrmann, H., Strelkov, S. V., Feja, B., Rogers, K. R., Brettel, M., Lustig, A., Haner, M., Parry, D. A., Steinert, P. M., Burkhard, P. & Aebi, U. (2000). The intermediate filament protein consensus motif of helix 2B: its atomic structure and contribution to assembly. *Journal of Molecular Biology* **298**, 817-832.

28. Steinert, P. M. & Parry, D. A. (1993). The conserved H1 domain of the type II keratin 1 chain plays an essential role in the alignment of nearest neighbor molecules in mouse and human keratin 1/keratin 10 intermediate filaments at the two- to four-molecule level of structure. *Journal of Biological Chemistry* **268**, 2878-2887.
29. Steinert, P. M. (1993). Structure, function, and dynamics of keratin intermediate filaments. *Journal of Investigatory Dermatology* **100**, 729-734.
30. Schaffeld, M., Herrmann, H., Schultess, J. & Markl, J. (2001). Vimentin and desmin of a cartilaginous fish, the shark scyliorhinus stellaris: sequence, expression patterns and *in vitro* assembly. *European Journal of Cell Biology* **80**, 692-702.
31. O'Shea, E. K., Klemm, J. D., Kim, P. S. & Alber, T. (1991). X-ray structure of the GCN4 leucine zipper, a two-stranded, parallel coiled coil. *Science* **254**, 539-544.
32. Lee, C. H., Kim, M. S., Chung, B. M., Leahy, D. J. & Coulombe, P. A. (2012). Structural basis for heteromeric assembly and perinuclear organization of keratin filaments. *Nature Structural and Molecular Biology* **19**, 707-715.
33. Janin, J., Bahadur, R. P. & Chakrabarti, P. (2008). Protein-protein interaction and quaternary structure. *Quarterly Reviews of Biophysics* **41**, 133-180.
34. Jones, S., Marin, A. & Thornton, J. M. (2000). Protein domain interfaces: characterization and comparison with oligomeric protein interfaces. *Protein Engineering* **13**, 77-82.
35. Powell, B. C., Cam, G. R., Fietz, M. J. & Rogers, G. E. (1986). Clustered arrangement of keratin intermediate filament genes. *Proceedings of the National Academy of Sciences of the United States of America* **83**, 5048-5052.
36. Bruce Fraser, R. D., Steinert, P. M. & Parry, D. A. D. (2003). Structural changes in trichocyte keratin intermediate filaments during keratinization. *Journal of Structural Biology* **142**, 266-271.
37. Marshall, R. C., Orwin, D. F. & Gillespie, J. M. (1991). Structure and biochemistry of mammalian hard keratin. *Electron Microscopy Reviews* **4**, 47-83.
38. Hatzfeld, M. & Weber, K. (1990). The coiled coil of *in vitro* assembled keratin filaments is a heterodimer of type I and II keratins: use of site-specific mutagenesis and recombinant protein expression. *Journal of Cell Biology* **110**, 1199-1210.
39. Robert, X. & Gouet, P. (2014). Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Research* **42**, 320-324.

40. Hatzfeld, M. & Weber, K. (1991). Modulation of keratin intermediate filament assembly by single amino acid exchanges in the consensus sequence at the C-terminal end of the rod domain. *Journal of Cell Science* **99**, 351-362.
41. Hatzfeld, M. & Weber, K. (1992). A synthetic peptide representing the consensus sequence motif at the carboxy-terminal end of the rod domain inhibits intermediate filament assembly and disassembles preformed filaments. *Journal of Cell Biology* **116**, 157-166.
42. Burkhard, P.; Stetefeld, J.; Strelkov, S. V. (2001). Coiled coils: a highly versatile protein folding motif. *Trends in cell biology* **11**, 82-88.
43. Lee, C.-H. & Coulombe, P. A. (2009). Self-organization of keratin intermediate filaments into cross-linked networks. *Journal of Cell Biology* **186**, 409-421.
44. Heitlinger, E., Peter, M., Haner, M., Lustig, A., Aebi, U. & Nigg, E. A. (1991). Expression of chicken lamin B2 in *E. coli*: characterization of its structure, assembly, and molecular interactions. *Journal of Cell Biology* **113**, 485-495.
45. Malavolta, L., Pinto, M. R. S., Cuvero, J. H. & Nakaie, C. R. (2006). Interpretation of the dissolution of insoluble peptide sequences based on the acid-base properties of the solvent. *Protein Science* **15**, 1476-1488.
46. Zhao, X., Pan, F. & Lu, J. R. (2008). Recent development of peptide self-assembly. *Progress in Natural Science* **18**, 653-660.
47. Wang, H., Parry, D. A. D., Jones, L. N., Idler, W. W., Marekov, L. N. & Steinert, P. M. (2000). *In vitro* assembly and structure of trichocyte keratin intermediate filaments: a novel role for stabilization by disulfide bonding. *The Journal of Cell Biology* **151**, 1459-1468.
48. Cui, H., Webber, M. J. & Stupp, S. I. (2010). Self-assembly of peptide amphiphiles: from molecules to nanostructures to biomaterials. *Biopolymers* **94**, 1-18.
49. Liu, H. & Bryson, W. G. (2002). A three-component model of the wool fibre-effects of morphology, elasticity and intermediate filament arrangement on fibre stiffness. *Journal of The Textile Institute* **93**, 121-131.
50. Ueli Aebi, W. E. F., Pamela Rew, and Tung-Tien Sun. (1983). The fibrillar substructure of keratin filaments unraveled. *Journal of Cell Biology* **97**, 1131-1143.
51. Feng, X. & Coulombe, P. A. (2015). A role for disulfide bonding in keratin intermediate filament organization and dynamics in skin keratinocytes. *Journal of Cell Biology* **209**, 59-72.

52. Judd, R. C. (2002). SDS-polyacrylamide gel electrophoresis of peptides. In the protein protocols handbook. *Humana Press.*

53. Ettinger, R. A., Liu, A. W., Nepom, G. T. & Kwok, W. W. (2000). Beta 57-Asp plays an essential role in the unique SDS stability of HLA-DQA1*0102/DQB1*0602 alpha beta protein dimer, the class II MHC allele associated with protection from insulin-dependent diabetes mellitus. *Journal of Immunology* **165**, 3232-3238.

54. Munson, G. W., Roher, A. E., Kuo, Y. M., Gilligan, S. M., Reardon, C. A., Getz, G. S. & LaDu, M. J. (2000). SDS-stable complex formation between native apolipoprotein E3 and β-amyloid peptides. *Biochemistry* **39**, 16119-16124.

55. Bromley, E. H., Sessions, R. B., Thomson, A. R. & Woolfson, D. N. (2009). Designed α-helical tectons for constructing multicomponent synthetic biological systems. *Journal of the American Chemical Society* **131**, 928-930.

56. Sun, T. T. & Green, H. (1978). Keratin filaments of cultured human epidermal cells. Formation of intermolecular disulfide bonds during terminal differentiation. *Journal of Biological Chemistry* **253**, 2053-2060.

57. Lorand, L., Parameswaran, K. N. & Murthy, S. N. P. (1998). A double-headed Gly-Pro-Arg-Pro ligand mimics the functions of the E domain of fibrin for promoting the end-to-end crosslinking of γ chains by factor XIII(a). *Proceedings of the National Academy of Sciences of the United States of America* **95**, 537-541.

58. Bolder, S. G., Vasbinder, A. J., Sagis, L. M. C. & van der Linden, E. (2007). Heat-induced whey protein isolate fibrils: conversion, hydrolysis, and disulphide bond formation. *International Dairy Journal* **17**, 846-853.

7 Chapter Seven

Summary, conclusions and suggestions for future work

7.1 Introduction

Self-assembling peptides have marked their niche amongst other biological molecules and biomolecular assemblies as molecular tectons with applications ranging from biotemplating, biomineralization for use in regenerative medicine, tissue engineering, 3D cell culture and biosensing devices [1]. The aims of this research work were to explore the protein-protein interfaces as a novel design source for self-assembling peptide tectons and to characterize the different self-assembled nanostructures formed by the peptide sequences. A large number of peptide sequences inspired from natural assemblies and biological sequences self-assemble into an array of different nanostructures[2]. Many such self-assembling peptide sequences have successfully found use as biomaterials, for example self-assembled peptide nanofibers and hydrogels in 3D cell culture as well as tissue engineering[3], responsive peptide gels and peptide based drug amphiphile for biomedical applications[4,5] and peptide nanowires bionanodevices[6,7]. Further, modifications in sequence and design strategy have allowed for the production of designer nanostructures such as honeycomb suitable for production of rationally designed bio-inspired surfaces and nanorings linked with functional biomolecules [8,9]. This approach provides a chance to exploit the natural self-assembly for the production of peptides that assemble in a precise manner in the 3-dimensional space[10,2].

To design peptide sequences from a bottom-up approach for use as molecular bio-based materials, it is important to explore new biological sources for the design of stable tectons and to understand their self-assembly process. Many self-assembling peptide sequences designed to date are largely limited to a set of biologically inspired sequences and designs[11,12,13,14,15]. Therefore, search for new design sources is warranted as they would add to the current library of self-assembling peptide sequences and provide for specific applications in bionanotechnology such as building blocks for design of conductive nanotubes/nanowires for nanocircuitry[16,17].

Molecular building blocks should have properties that make them suitable for downstream applications including responsive hydrogels for biomedical sciences, biotemplating, sensing, and conductive nanowires, and also to have better control over the self-assembly process. An important property the building blocks should possess foremost is the ability to self-assemble independently into well-ordered and defined forms. This is the basic feature to develop general principles of self-assembly into hierarchical structures from simple peptide sequences[18].

The second important feature is that the building blocks should have a specific chemistry that can be explored for control over the self-assembly process. The chemistry of the peptide sequence defines the behaviour of the peptide in response to a physiological trigger such as concentration, hydrophobicity, pH, and ionic strength. Specific sequence modifications to existing self-assembling peptide sequences have allowed for growth and fabrication of peptide assemblies on surface[19] and self-assembling peptide linked to metal nanoparticles[20,21].

Peptides offer a huge advantage of customization through attachment points for decoration and chemical modifications, thus providing unique functionality to the peptide nanostructures. Certain peptide sequences that are made up of two or more blocks such as the multi-domain peptides containing charged residues, chemically linked to an active peptide sequence self-assembled into extended fibrillar assemblies with the biologically functional sequence expressed on the surface[22].

Peptide sequences containing an extended histidine sequence due to their affinity for metal ions such as Ni^{2+} could be functionalized with gold nanoparticles[23]. Peptide functionalized with gold nanoparticles can provide a range of surface functionalities and the peptide element provides a rich resource to control structures and properties of gold particles. This is particularly attractive towards creating probes for sensing molecules including ions, minerals and biomolecules as well as excellent models as drug delivery vehicles.

Thirdly, it is very important that the self-assembled peptide structure is stable, especially from the perspective of a viable downstream application such as scaffolds for templating synthetic and biological polymers, nanofabrication and nanolithography. The availability of peptide tectons those are tolerant to a range of pH, ionic strength and exhibit reversible self-assembly are desirable for their fabrication and incorporation into bio-nanodevices. Additionally, peptide sequences tolerant to small sequence modifications provide an added advantage to explore as bio-organic building blocks that are responsive to change in pH, redox state, and ionic strength. These properties are particularly interesting for the use of peptide based material as biomaterial in regenerative medicine[3] and to build nanodevices[7] .

Interface sequences within different protein oligomers that exist in complementary non-covalent close contacts were identified to meet these requirements. Protein interfaces are evolutionarily optimized to drive self-assembly of protein units in a controlled and regulated manner[24,25]. This property makes them attractive as sources of peptide sequences that can self-assemble while retaining some properties of the parent interface.

Although many self-assembling peptide sequences designed from biological motifs are reported, protein-protein interfaces, more specifically protein-β interfaces, and coiled-coil interface as the source for self-assembling peptide tectons is investigated here for the first time. This thesis aimed to investigate the self-assembly of short peptide sequences designed from interfaces of oligomer proteins, as a molecular biomaterial. A bottom-up approach was undertaken to design peptide sequences from protein-protein interfaces, naturally fine-tuned for self-assembly.

The self-assembly of the peptide sequences in solution was characterized using CD spectroscopy and ThT binding assays. The structure and morphology of the peptide assemblies were determined using biophysical techniques like SAXS, FTIR, and TEM. The β-sheet type self-assembling peptide sequences were further used to template the higher order nanotubular assembly of *h*Prx3 protein.

7.2 Self-assembled β-sheet type peptide nanostructures

The first part of the design strategy focussed on the simple β-continuous protein-protein interfaces reported in nearly 15% of protein homo-oligomers[26]. A PDB-wide search for protein homo-oligomers focused on simple β-continuous interfaces led to the selection of four different oligomer protein sources. Such interfaces within protein homo-oligomers have not been previously described as a source for peptide building blocks that could self-assemble in isolation.

The peptide sequences exhibited intrinsic ability to form β-sheets in a solution that formed birefringent hydrogels. EM showed that the peptide sequences self-assembled in solution to form extended nanostructures. SAXS showed that the molecular organization of β-sheets within the extended peptide nanostructures was consistent with those previously described for peptide hormones[27] and peptide sequences that self-assemble into amyloid-like fibrils[28]. The liquid crystallinity of the peptide hydrogels that is described for reversibly self-assembling peptide sequences[29], however, makes them different to the insoluble amyloid aggregates. This property makes them highly attractive as responsive self-assembling building blocks. ThT binding assays also led to predict self-assembly models for the peptide sequences that was largely a single state model, with the only exception of DAPDC peptide.

Conclusions derived from this work:

(I) Peptide sequences derived from protein-β interfaces of unrelated protein oligomers self-assemble in solution into extended nanostructures in a reversible manner.

(II) The DAPDC peptide derived from the homo-tetramer interface of *M. tuberculosis* DAPDC protein showed a higher critical assembly concentration compared to the other peptides.

The results show the inherent ability of diverse peptide sequences designed from unrelated protein oligomers to self-assemble in solution, with properties arising from the parent interface motifs[30]. This work adds to the current repertoire of peptide tectons and also show the potential of the peptide sequences as stable building blocks for nanofabrication and functionalization.

7.3 Prx3 peptide self-assembly is tolerant to specific sequence variations

A mutational approach was used as a tool to understand the self-assembly of the Prx3 peptide through judiciously varying the histidine residue in the peptide sequence. The introduction of a subtle change in the chemistry of the peptide sequence is a quick and easy way to identify key interactions necessary for self-assembly[31,32]. The variant peptide sequences exhibited self-assembling properties noted for the parent Prx3 peptide, consistent with β-sheet propensity. ThT binding studies showed that the self-assembly of peptide sequence into extended β-sheets is tolerant to the specific change of the histidine residue. ThT assays also highlighted that the specific sequence change of histidine to tyrosine and the larger tryptophan residue contributed to the higher critical assembly concentration of the parent Prx3 peptide.

EM and SAXS were useful in stating that the variant peptide sequences self-assembled in solution into extended nanostructures that exhibited the generic architecture of β-sheets organized in a repetitive manner over a long distance, consistent with amyloid-like fibrils[33]. EM also revealed that the H3Y peptide variant self-assembled into polymorphic nanostructures ranging from nanotapes to wide peptide nanosheets favoured at high concentration of the peptide in water. Polymorphism is a generic feature of amyloid-like fibrils[34]. Such assembly, however, is also reported to occur through lateral association of individual extended amyloid-like fibrils into wider tapes as a consequence of the compaction of the nanostructures adsorbed, and air dried onto the EM grids[35].

This property of the peptide sequence is attractive to be exploited for production and subsequent fabrication of peptide nanosheets to create bio-inspired surfaces for nanofabrication. Characterization of *h*Prx3 muteins containing the mutations to the peptide sequence at the homodimer protein-β interface suggested a correlation in the behaviour of the variant peptide motif within the 3-dimensional protein fold and the isolated peptide sequence.

Conclusions derived from this work:

(I) Hydrophobic variations introduced in the Prx3 peptide through replacement of histidine residue does not substantially affect the self-assembly of the peptide sequence.

(II) The specific sequence change of the Prx3 peptide by replacement of histidine with aromatic tyrosine (H3Y) and the larger tryptophan group (H3W) shows a higher critical assembly concentration compared to the other variations.

(III) The specific mutation of the histidine residue to tryptophan within the parent protein (H136W) weakened the protein β-interface.

7.4 Self-assembling peptides template higher order assembly of *h*Prx3

In an attempt to produce supramolecular protein assemblies, the strategy mixing of self-assembled Prx3 peptide nanotapes with *h*Prx3 protein toroids was trailed. This strategy interestingly led to the formation of micron-long nanotubes of *h*Prx3 protein toroids templated through extended peptide nanotapes. Such extended assembly of the protein is described here for the first time. The strategy used to produce the extended protein nanostructures is described in figure 7.1.

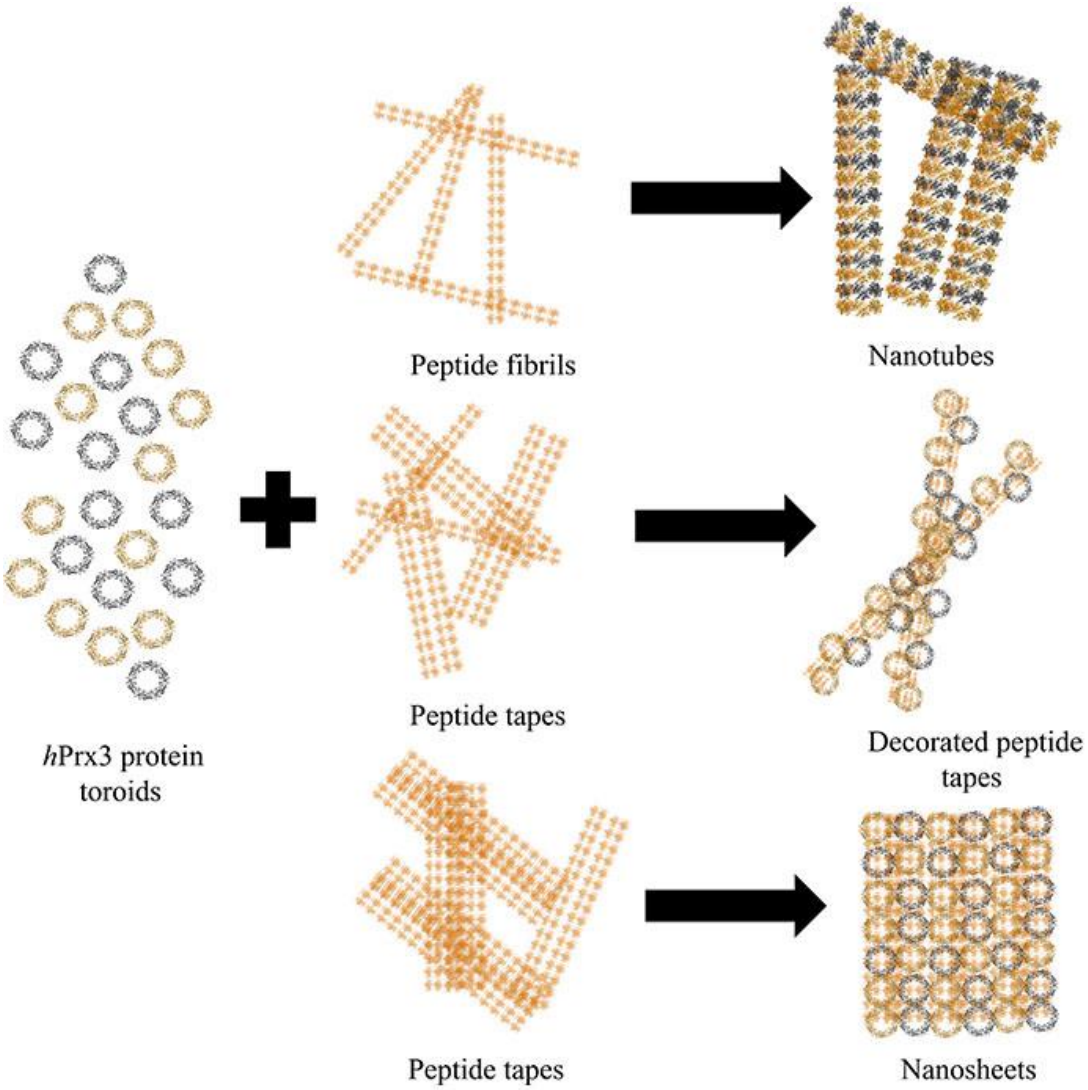


Figure 7.1: *Schematic diagram of the strategy for production of different extended nanostructures of hPrx3 protein (PDB: 5JCG) using peptide nanostructures. Three different outcomes were achieved in this thesis work by mixing peptide nanostructures with hPrx3 protein toroids: (i) Extended nanotubes, (ii) Peptide tapes decorated with protein toroids and (iii) Extended nanosheets.*

The specific assembly of the *h*Prx3 toroids is induced by all the self-assembling β-sheet type peptide sequences described in this thesis work. Furthermore, mixing of wide extended nanotapes of the H3Y variant peptide with *h*Prx3 led to the formation of decorated peptide tapes and protein nanosheets.

Conclusions derived from this work:

(I) Self-assembling peptide sequences template higher order assembly of *h*Prx3 protein toroids into extended micron long nanotubes, peptide tapes decorated with toroids and protein nanosheets.

(II) The self-assembly of the extended peptide nanostructures with the *h*Prx3 protein toroids is of electrostatic nature.

(III) Varying ratio of peptide mixed with *h*Prx3 protein toroids allow some control on the overall size of nanotubular structures.

The results highlighted the ability of self-assembling peptide nanostructures with diverse chemistry to template complex protein molecules in solution into an ordered array of nanostructures. The extended protein nanotubes show potential to be used for the creation of functionalized nanostructures that can be fabricated into bio-nanodevices.

7.5 Self-assembled helical coiled coil peptide nanostructures

In another approach to design responsive self-assembling peptide tectons, the coiled-coil interface of *ovine* IF protein keratin was investigated. IF protein is rich in coiled-coil motifs and self-assemble into extended filaments[36] and fibres[37] that show good mechanical strength and are biodegradable. The hallmark of many such coiled coil assemblies is the organization of the amino acid residues as seven residue heptad sequence, with a specific distribution of hydrophobic and hydrophilic residues. This amphipathic design is the basis of the diverse set of nanostructures formed by heptad based sequences[38]. In the specific case of IF hard α-keratins, the central coiled coil region is richly distributed with cysteine residues that make them redox sensitive.

Peptide sequences designed from the consensus sequence motif of type I and type II keratins self-assembled in solution into extended filament-like nanostructures. The dimensions and appearance of the formed nanostructures closely resembled the IFs formed by full-length keratins that self-assemble as obligate heteromers. The extended filament assembly of the specific peptide sequences has been described here for the first time. SDS-PAGE suggested that the peptide strands within the extended nanostructures be potentially involved in disulphide linking.

Conclusions derived from this work:

(I) Peptide sequences derived from the consensus sequence motif of *ovine* IF keratins self-assembled in solution with extended filament-like nanostructures observed for K81P1 and K31P2 peptides.

(II) The extent of self-assembly of the peptide sequences into extended filaments is favoured under oxidising conditions and acidic pH.

(III) The self-assembly of the peptide sequences exhibits some sensitivity to the change in the redox state of the peptide.

The results highlight coiled-coil interface of *ovine* keratins as a source of peptide tectons.

7.6 Suggestions for future work

In this thesis work, self-assembly of peptide sequences derived from protein-β interfaces and coiled-coil protein interface was described. Despite promising results, it was difficult to crystallize these sequences and model the atomic level structural arrangements of the peptide strands within the extended nanostructures. EM showed that the peptides self-assembled into extended nanostructures with polymorphism observed in the case of Prx3 peptide and a variant of the peptide H3Y. Polymorphism is an important feature of amyloid fibrils formed by diverse

peptide sequences and proteins. In contrast, it has also been noted that the wide tape like assemblies observed by EM is a consequence of air drying of the sample on EM grid that leads to the lateral association of the peptide fibrils into tapes or ribbons that coil[35]. Future work should, therefore, be focused on the in-solution analysis of the extended peptide nanostructures through cryo-EM. It will allow a good comparison between the morphology of the peptide nanostructures in solution and that observed by TEM.

3-D reconstruction of the cryo-EM micrographs developed using single particle analysis can provide insight into the arrangement of the β-sheets within the extended peptide nanostructures. MAS NMR coupled with cryo-EM have been successful in generating atomic resolution structures of amyloid-like fibrils and provide a detailed structural arrangement of β-sheets within the extended amyloid fibrils[39]. This technique can further allow gazing on the different arrangements of the β-sheet in the polymorphic nanotapes of H3Y peptide variant. Molecular dynamic simulation studies of the Prx3 parent peptide and the hydrophobic peptide variants can provide detailed insight into the arrangement of the β-strands and sheets in the peptide tapes and the molecular interactions involved in the self-assembly.

Crystallization trials performed with an aim to understand the molecular model of the self-assembled peptide β-sheet did not show any positive outcome, for the conditions trailed. It is difficult to crystallize peptide sequences that rapidly self-assemble into extended amyloid-like fibrils. Only a few peptide sequences have been crystallized thus far[40]. One important factor that hinders in the crystallization of the peptide sequences is the change in solute concentration. Solute effects are known to alter the dynamic equilibrium between the β-sheets and aggregates, and increase in solute concentration accelerate the aggregation of the peptide[41]. ThT binding assays suggested a higher critical assembly concentration for the DAPDC peptide, and aromatic

variants of Prx3 peptide: H3W and H3Y respectively. These peptide sequences would prove to be beneficial for the purpose of crystallization. Therefore future work can be directed towards screening for conditions that favor crystallization of these peptides sequences.

Self-assembled peptide nanostructures were used to template extended nanotubes and nanosheets of *h*Prx3 protein toroids. Despite DLS suggesting the presence of large extended structures in solution however, it was difficult to detect the extended peptide-protein nanotubular assembly by AUC. Also, it was difficult to characterize the extended nanotubular structures observed by TEM for peptide-protein mix containing a large molar excess of peptide mixed with protein by DLS or AUC due to the increase in macroscopic viscosity of the sample[42]. Future work in this direction should focus on the in solution cryo-EM analysis of the viscous peptide-protein mixes to characterize the extended nanotubular structures observed by TEM. This will help to address, whether the extended protein-peptide assembly exists in solution, or the templating of the *h*Prx3 toroids onto peptide nanostructures is a consequence of air-drying of the sample onto EM grid due to macromolecular crowding[43].

TEM showed that the peptide-protein assembly is specific to *h*Prx3, and the interaction between the peptide nanostructures and *h*Prx3 protein toroids is of electrostatic nature. However, it was difficult to describe the role of specific residues or regions within the *h*Prx3 protein toroid that are involved in the specific association of the protein toroids to the extended peptide nanostructures. 3-D reconstructions of the cryo-EM micrographs would give insight into the different associated states of the peptide nanostructures with *h*Prx3 protein toroids as observed with the H3Y variant peptide. Modeling of the different associated states with the crystal structure of extended stacks of *h*Prx3 toroids[44] would help to predict the specific regions of the *h*Prx3 protein toroid involved in the interaction with the extended peptide nanostructures.

Mutations specific to these identified regions would further allow targeting of the specific molecular interactions that drive the assembly of the *h*Prx3 protein toroids to the peptide nanostructures.

Peptide sequences derived from the coiled-coil interface of *ovine* keratin protein self-assembled into extended filament-like nanostructures. The peptide sequences were designed incorporated with cysteine residues with an aim to develop responsive peptide tectons. SDS-PAGE suggested a potential role of disulfide linking in peptide self-assembly, although it was difficult to differentiate disulfide-linked peptide aggregates from SDS-resistant species. Mass spectrometry (MS) provides a potent tool to differentiate between the different disulfide cross-linked forms of the peptide and the resistant species[45]. Future work should, therefore, involve MS analysis of the peptide sequences under oxidizing/non-reducing and reducing conditions to characterize the different peptide aggregates detected by SDS-PAGE.

The self-assembling peptide sequences described here have a diverse surface chemistry that can be exploited for the development of functional biomaterial. One such approach was trailed through judicious modification of the Prx3 peptide sequence, which led to the development of a prototype model of functional organic semiconductor (Appendix II).

7.7 Final conclusion

To conclude, this thesis illustrates new sources for the design of peptide tectons that self-assemble into well-ordered nanostructures in solution. Moreover, practical approaches for characterization of peptide self-assembly *in vitro* and high-resolution techniques that provided structural organization of the peptide within the self-assembled nanostructures are presented.

7.8 References

1. Jaime Castillo, W. E. S., Luigi Sasso. (2012). Self-assembled peptide nanostructures: advances and applications in nanotechnology. *Pan Standard*.
2. Mendes, A. C., Baran, E. T., Reis, R. L. & Azevedo, H. S. (2013). Self-assembly in nature: using the principles of nature to create complex nanobiomaterials. *Wiley Interdisciplinary Reviews: Nanomedicine and Nanobiotechnolgy* **5**, 582-612.
3. Koutsopoulos, S. (2016). Self-assembling peptide nanofiber hydrogels in tissue engineering and regenerative medicine: Progress, design guidelines, and applications. *Journal of Biomedical Materials Research Part A* **104**, 1002-1016.
4. Rad-Malekshahi, M., Lempsink, L., Amidi, M., Hennink, W. E. & Mastrobattista, E. (2016). Biomedical spplications of self-assembling peptides. *Bioconjugate Chemistry* **27**, 3-18.
5. Liu, J. & Zhao, X. (2011). Design of self-assembling peptides and their biomedical applications. *Nanomedicine* **6**, 1621-1643.
6. Domigan, L. J. (2013). Proteins and peptides as biological nanowires: towards biosensing devices. *Methods in Molecular Biology* **996**, 131-152.
7. Lakshmanan, A., Zhang, S. & Hauser, C. A. (2012). Short self-assembling peptides as building blocks for modern nanodevices. *Trends in Biotechnol*ogy**30**, 155-165.
8. Abb, S., Harnau, L., Gutzler, R., Rauschenbach, S. & Kern, K. (2016). Two-dimensional honeycomb network through sequence-controlled self-assembly of oligopeptides. *Nature Communications* **7**, 1-7.
9. Park, S., Lee, M. & Shin, S. (2015). A molecular dynamics study on controlling the self-assembly of β-sheet peptides with designer nanorings. *Chemistry: An Asian Journal* **10**, 1684-1689.
10. Zhang, S. (2003). Building from the bottom up. *Materials Today* **6**, 20-27.
11. Ghadiri, M. R., Granja, J. R., Milligan, R. A., McRee, D. E. & Khazanovich, N. (1993). Self-assembling organic nanotubes based on a cyclic peptide architecture. *Nature* **366**, 324-327.
12. Lu, K., Jacob, J., Thiyagarajan, P., Conticello, V. P. & Lynn, D. G. (2003). Exploiting amyloid fibril lamination for nanotube self-assembly. *Journal of the American Chemical Society* **125**, 6391-6393.

13. Valery, C., Paternostre, M., Robert, B., Gulik-Krzywicki, T., Narayanan, T., Dedieu, J. C., Keller, G., Torres, M. L., Cherif-Cheikh, R., Calvo, P. & Artzner, F. (2003). Biomimetic organization: Octapeptide self-assembly into nanotubes of viral capsid-like dimension. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 10258-62.
14. Hamley, I. W. (2011). Self-assembly of amphiphilic peptides. *Soft Matter* **7**, 4122-4138.
15. Woolfson, D. N., Bartlett, G. J., Bruning, M. & Thomson, A. R. (2012). New currency for old rope: From coiled-coil assemblies to α-helical barrels. *Current Opinion in Structural Biology* **22**, 432-441.
16. Valery, C., Artzner, F. & Paternostre, M. (2011). Peptide nanotubes: molecular organisations, self-assembly mechanisms and applications. *Soft Matter* **7**, 9583-9594.
17. Rajagopal, K. & Schneider, J. P. (2004). Self-assembling peptides and proteins for nanotechnological applications. *Current Opinion in Structural Biology* **14**, 480-486.
18. Mandal, D., Nasrolahi Shirazi, A. & Parang, K. (2014). Self-assembly of peptides to nanostructures. *Organic & Biomolecular Chemistry* **12**, 3544-3561.
19. Lee, J. S., Ryu, J. & Park, C. B. (2009). Bio-inspired fabrication of superhydrophobic surfaces through peptide self-assembly. *Soft Matter* **5**, 2717-2720.
20. Kasotakis, E. & Mitraki, A. (2013). Designed self-assembling peptides as templates for the synthesis of metal nanoparticles. *Methods in Molecular Biology* **996**, 195-202.
21. Yan, J., Pan, Y., Cheetham, A. G., Lin, Y.-A., Wang, W., Cui, H. & Liu, C.-J. (2013). One-step fabrication of self-assembled peptide thin films with highly dispersed noble metal nanoparticles. *Langmuir* **29**, 16051-16057.
22. Aulisa, L., Dong, H. & Hartgerink, J. D. (2009). Self-Assembly of multidomain peptides: Sequence variation allows control over cross-linking and viscoelasticity. *Biomacromolecules* **10**, 2694-2698.
23. Hamley, I. W., Kirkham, S., Dehsorkhi, A., Castelletto, V., Adamcik, J., Mezzenga, R., Ruokolainen, J., Mazzuca, C., Gatto, E., Venanzi, M., Placidi, E., Bilalis, P. & Iatrou, H. (2014). Self-assembly of a model peptide incorporating a hexa-histidine sequence attached to an oligo-alanine sequence, and binding to gold NTA/nickel nanoparticles. *Biomacromolecules* **15**, 3412-3420.

24. Tsai, C. J., Lin, S. L., Wolfson, H. J. & Nussinov, R. (1996). Protein-protein interfaces: architectures and interactions in protein-protein interfaces and in protein cores. Their similarities and differences. *Critical Reviews in Biochemistry and Molecular Biology* **31**, 127-152.

25. Dey, S., Pal, A., Chakrabarti, P. & Janin, J. (2010). The subunit interfaces of weakly associated homodimeric proteins. *Journal of Molecular Biology* **398**, 146-160.

26. Nooren, I. M. A. & Thornton, J. M. (2003). Structural characterisation and functional significance of transient protein-protein interactions. *Journal of Molecular Biology* **325**, 991-1018.

27. van Grondelle, W., Iglesias, C. L., Coll, E., Artzner, F., Paternostre, M., Lacombe, F., Cardus, M., Martinez, G., Montes, M., Cherif-Cheikh, R. & Valery, C. (2007). Spontaneous fibrillation of the native neuropeptide hormone Somatostatin-14. *Journal of Structural Biology* **160**, 211-223.

28. Morris, K. L. & Serpell, L. C. (2012). X-ray fibre diffraction studies of amyloid fibrils. *Methods in Molecular Biology* **849**, 121-135.

29. Valery, C., Artzner, F., Robert, B., Gulick, T., Keller, G., Grabielle-Madelmont, C., Torres, M. L., Cherif-Cheikh, R. & Paternostre, M. (2004). Self-association process of a peptide in solution: from β-sheet filaments to large embedded nanotubes. *Biophysical Journal* **86**, 2484-2501.

30. Valery, C., Pandey, R. & Gerrard, J. A. (2013). Protein β-interfaces as a generic source of native peptide tectons. *Chemical Communications* **49**, 2825-2827.

31. Valery, C., Pouget, E., Pandit, A., Verbavatz, J. M., Bordes, L., Boisde, I., Cherif-Cheikh, R., Artzner, F. & Paternostre, M. (2008). Molecular origin of the self-assembly of lanreotide into nanotubes: a mutational approach. *Biophysical Journal* **94**, 1782-1795.

32. Mannini, B., Mulvihill, E., Sgromo, C., Cascella, R., Khodarahmi, R., Ramazzotti, M., Dobson, C. M., Cecchi, C. & Chiti, F. (2014). Toxicity of Protein Oligomers Is Rationalized by a Function Combining Size and Surface Hydrophobicity. *ACS Chemical Biology* **9**, 2309-2317.

33. Langkilde, A. E., Morris, K. L., Serpell, L. C., Svergun, D. I. & Vestergaard, B. (2015). The architecture of amyloid-like peptide fibrils revealed by X-ray scattering, diffraction and electron microscopy. *Acta Crystallographica Section D Biological Crystallography* **71**, 882-895.

34. Fändrich, M., Meinhardt, J. & Grigorieff, N. (2009). Structural polymorphism of Alzheimer Aβ and other amyloid Fibrils. *Prion* **3**, 89-93.

35. Gosal, W. S., Clark, A. H., Pudney, P. D. A. & Ross-Murphy, S. B. (2002). Novel amyloid fibrillar networks derived from a globular protein: β-Lactoglobulin. *Langmuir* **18**, 7174-7181.

36. Harland, D. P., Caldwell, J. P., Woods, J. L., Walls, R. J. & Bryson, W. G. (2011). Arrangement of trichokeratin intermediate filaments and matrix in the cortex of Merino wool. *Journal of Structural Biology* **173**, 29-37.

37. Claudio Tonin, A. A., Alessio Varesano and Claudia Vineis. (2010). Keratin-based nanofibres: In nanofibers. *InTech*.

38. Woolfson, D. N. (2010). Building fibrous biomaterials from α-helical and collagen-like coiled-coil peptides. *Biopolymers* **94**, 118-127.

39. Fitzpatrick, A. W. P., Debelouchina, G. T., Bayro, M. J., Clare, D. K., Caporini, M. A., Bajaj, V. S., Jaroniec, C. P., Wang, L., Ladizhansky, V., Müller, S. A., MacPhee, C. E., Waudby, C. A., Mott, H. R., De Simone, A., Knowles, T. P. J., Saibil, H. R., Vendruscolo, M., Orlova, E. V., Griffin, R. G. & Dobson, C. M. (2013). Atomic structure and hierarchical assembly of a cross-β amyloid fibril. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 5468-5473

40. Sawaya, M. R., Sambashivan, S., Nelson, R., Ivanova, M. I., Sievers, S. A., Apostol, M. I., Thompson, M. J., Balbirnie, M., Wiltzius, J. J., McFarlane, H. T., Madsen, A. O., Riekel, C. & Eisenberg, D. (2007). Atomic structures of amyloid cross-β spines reveal varied steric zippers. *Nature* **447**, 453-457.

41. Schmit, J. D., Ghosh, K. & Dill, K. (2011). What drives amyloid molecules to assemble into oligomers and fibrils? *Biophysical Journal* **100**, 450-458.

42. Wieczorek, G. & Zielenkiewicz, P. (2008). Influence of macromolecular crowding on protein-protein association rates—a brownian dynamics study. *Biophysical Journal* **95**, 5030-5036.

43. Hu, Z., Jiang, J. & Rajagopalan, R. (2007). Effects of macromolecular crowding on biochemical reaction equilibria: a molecular thermodynamic perspective. *Biophysical Journal* **93**, 1464-1473.

44. Yewdall, N. A., Venugopal, H., Desfosses, A., Abrishami, V., Yosaatmadja, Y., Hampton, M. B., Gerrard, J. A., Goldstone, D. C., Mitra, A. K. & Radjainia, M. (2016). Structures of human peroxiredoxin 3 suggest self-chaperoning assembly that maintains catalytic State. *structure* **24**, 1120-1129.

45. Xu, H., Zhang, L. & Freitas, M. A. (2008). Identification and characterization of disulfide bonds in proteins and peptides from tandem MS data by use of the massmatrix MS/MS search engine. *Journal of Proteome Research* **7**, 138-144.

Appendix I

A1.1 Supplementary figures for chapter 5

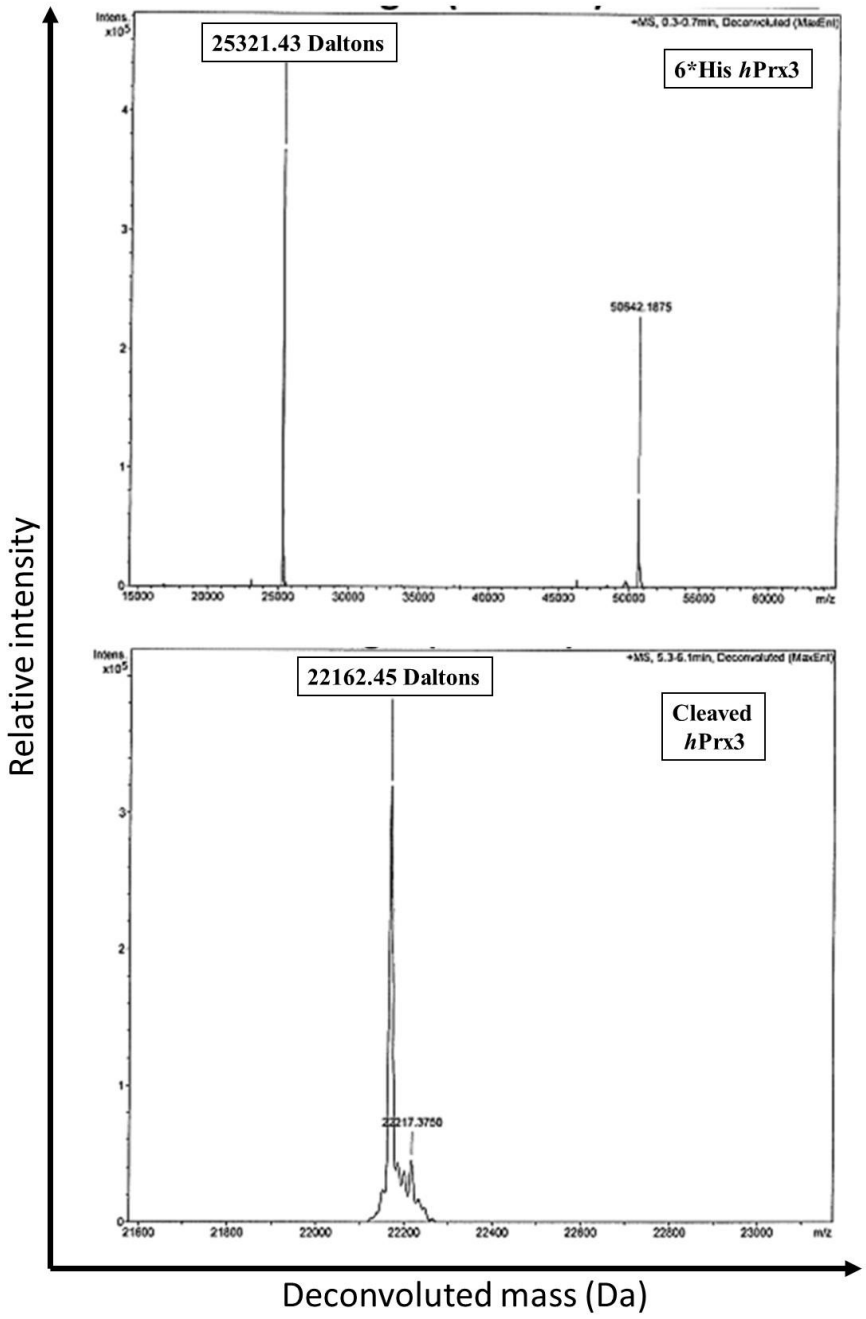


Figure A1: *Deconvoluted mass spectrum of tagged 6*His hPrx3 and cleaved hPrx3 (tag removed): The difference in the mass of the cleaved protein confirms cleavage and absence of any band corresponding to the ~ 3kDa peptide His-tag confirmed a pure sample of untagged hPrx3. The m/z scan range is 300 to 2900.*

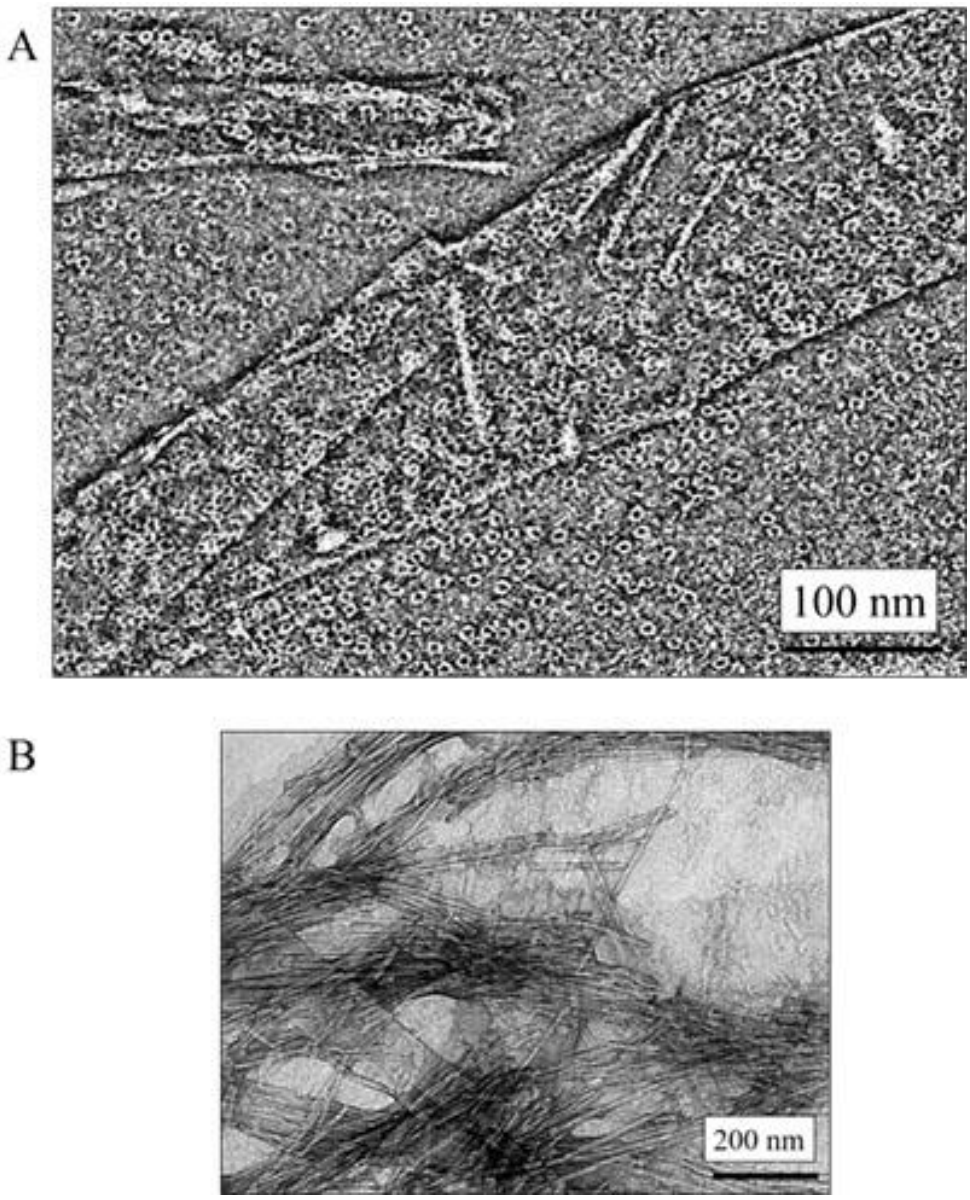


Figure A2: *TEM of peptide-protein mix in reducing pH 8.0 buffer with 150 mM NaCl dialyzed into high salt pH 8.0 buffer overnight, showing mostly free peptide fibrils, short nanotubes, and free protein toroids. The arrows point to the protein toroids associated with peptide-fibrils in a non-specific manner. (I) Peptide control sample at 5 mg/mL shows extended nanostructures and (II) Control protein sample concentration at 0.1 mg/mL shows mostly protein toroids. (Scale bar is 100 nm)*

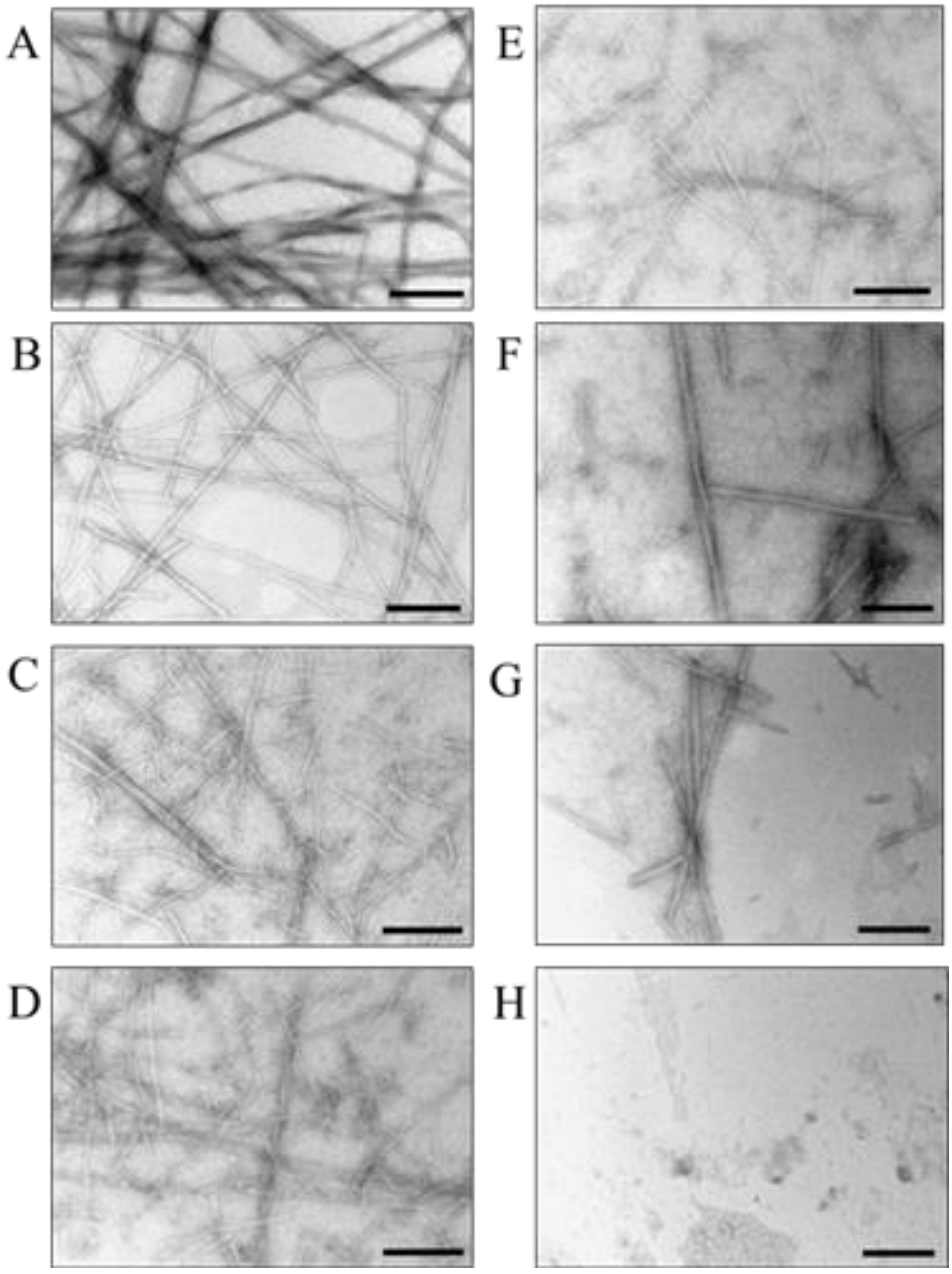


Figure A3: *Representative TEM micrographs of control Prx3 peptide acetate in reducing pH 8.0 buffer at different concentrations. The peptide acetate concentration is: (A)5mg/mL, (B)2.5 mg/mL, (C)1 mg/mL, (D)0.5 mg/mL, (E)0.17 mg/mL, (F)17 µg/mL, (G)0.34 µg/mL (equimolar concentration to protein) and (H) 6.8 ng/mL (below equimolar concentration to protein). (Scale bar is 200 nm)*

Appendix II

A2.1 Collaboration work done with Hodgkiss group

Functional organic semiconductor functionalised through self-assembling peptides:

This work was done in collaboration with the Hodgkiss group at Victoria University, Wellington (N.Z.). With an aim to develop self-assembling peptide-polymer conjugates, variants of the Prx3 peptide were designed with composition as detailed in table 2.1.

Table 2.1: *Variant peptide sequences*

Peptide	Composition	Comments
Prx3	Ac-IKHLSVN-NH_2	*Bovine* peroxiredoxin 3
V1	Ac-IRHLSVN	
V2	-GIRHLSVN-	Variants of Prx3 peptide
V3	-GIRHLSVNEEE-	
V4	Ac-EEEIRHLSVN-Et-NH_2	

Note: The variant peptide sequences were supplied as lyophilized powder with purity of 95% as judged by NMR.

Characterization of the variant peptides in solution:

Far UV circular dichroism spectra of the variant peptide sequences in water at 0.01 mg/mL (Section 2.6.4) showed that the synthetic variant peptides suggested that the peptide sequences show a spectral shape that is characteristic that for β-sheets in the native Prx3 peptide sequence (Figure A2.1). The formation of β-sheets by the variant peptide sequences was further confirmed by ThT-binding assays (Section 2.6.6), for peptide samples in the 1-10 mg/mL range, equilibrated for 24 hours at 4°C (Figure A2.2).

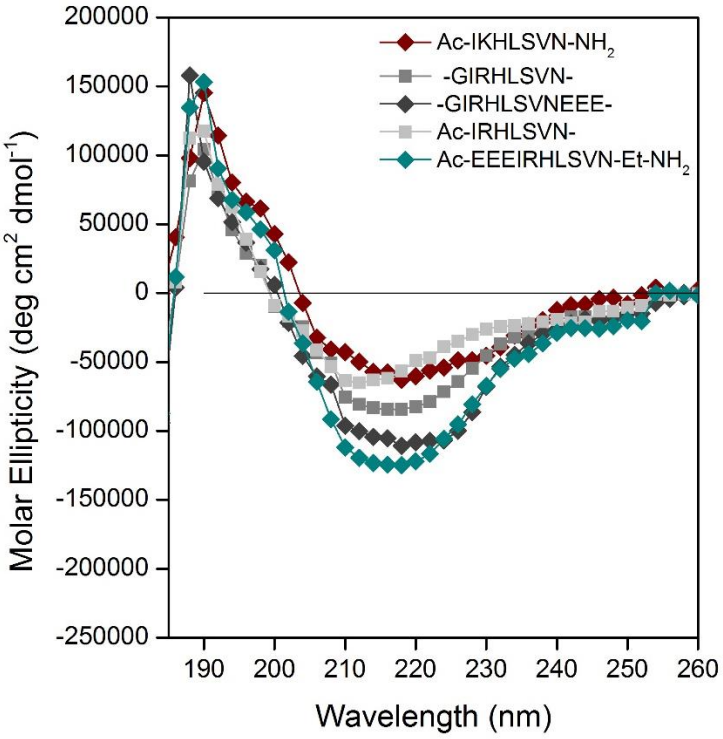


Figure A2.1: *Far UV circular dichroism spectra of Prx3 peptide and variant peptide acetates in water at 0.01 mg/mL. The peptide acetate in water samples was equilibrated for 24 hours at 4°C and degassed before analysis.*

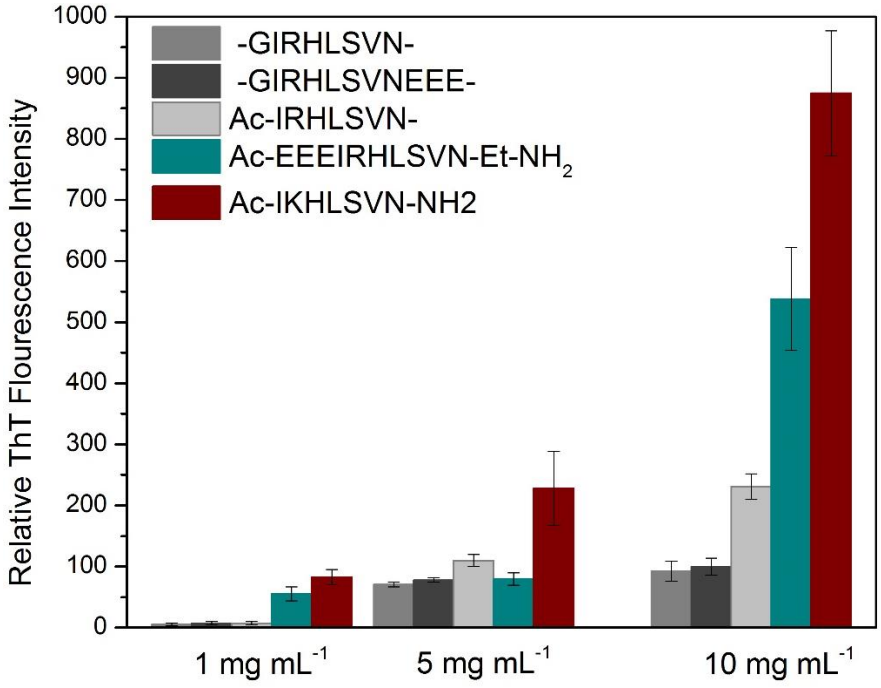


Figure A2.2: *ThT binding assay by dilution for Prx3 peptide and variant peptide acetates in water equilibrated for 24 hours at 4°C.*

Characterization of the variant peptides by TEM:

The variant peptide acetates in water at 1-5 mg/mL analysed by TEM (Section 2.6.8) showed that the variant peptide sequences self-assembled in water into nanostructures with extended morphology. Variant peptide sequence Ac-IRHLSVN however, was the only exception with the absence of any extended self-assembly (Figure A2.3).

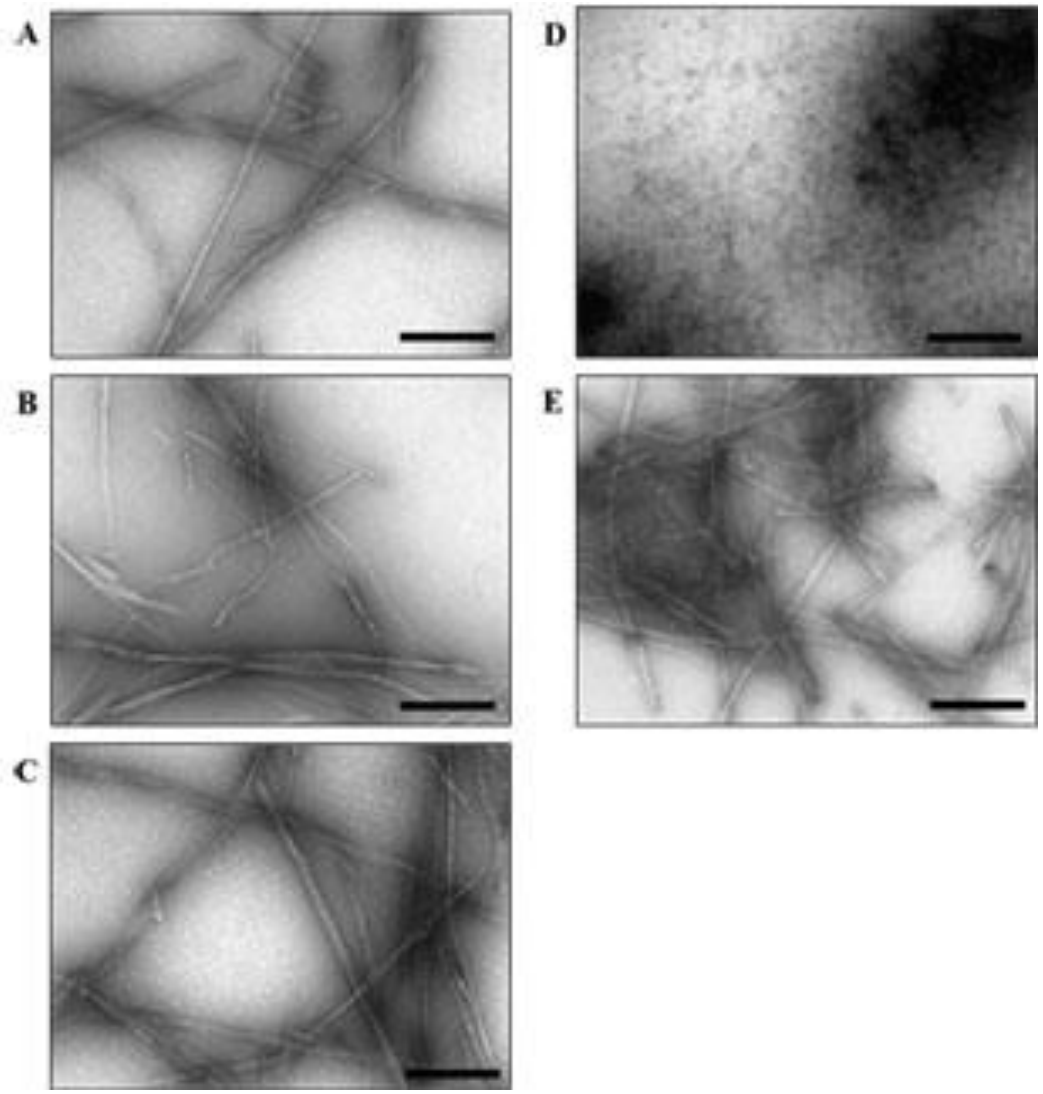


Figure A2.3: *Representative TEM micrographs of peptide acetates in water equilibrated for 24 hours at 4°C: (A) Ac-EEEIRHLSVN-ethyl-NH_2, (B) –GIRHLSVNEEE-, (C) –GIRHLSVN-, (D) Ac-IRHLSVN-, (E) Ac-IKHLSVN-NH_2. (Scale bar is 200 nm)*

The sequences described here were further used to design perylene diimide-peptide conjugates that were used to fabricate functional semiconductor prototype. The details of this work have been published in Eakins *et al.* (2015).